ALIBABA CLOUD

阿里云

E-MapReduce

集群管理

文档版本：20220615

阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。

2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。

3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。

4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。

5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。

6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	危险 重置操作将丢失用户配置数据。
警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	警告 重启操作将导致业务中断，恢复业务时间约十分钟。
注意	用于警示信息、补充说明等，是用户必须了解的内容。	注意 权重设置为0，该服务器不会再接受新请求。
说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击**设置**> **网络**> **设置网络类型**。
粗体	表示按键、菜单、页面名称等UI元素。	在**结果确认**页面，单击**确定**。
Courier字体	命令或代码。	执行 `cd /d C:/window` 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	`bae log list --instanceid` *Instance_ID*
[] 或者 [a|b]	表示可选项，至多选择一个。	`ipconfig [-all|-t]`
{} 或者 {a|b}	表示必选项，至多选择一个。	`switch {active|stand}`

目录

1.集群类型

本文为您介绍E-MapReduce支持的集群类型以及各集群相关的操作。

介绍

集群	描述	重要操作
Hadoop	• 提供半托管的Hadoop、Hive和Spark离线大规模分布式数据存储和计算。 • 提供Presto和Impala交互式查询。 • 提供Oozie等Hadoop生态圈的组件。	• 创建集群 • 登录集群 • 释放集群
Zookeeper	提供独立的分布式一致性锁服务，适用于大规模的Hadoop集群、HBase集群和Kafka集群。	概述

DataScience	主要面向大数据+AI场景，提供Hive和Spark离线大数据ETL和TensorFlow模型训练，您可以选择CPU+GPU的异构计算框架，通过英伟达GPU对部分深度学习算法进行高性能计算。	• 概述
		• 创建集群
		• 查看服务信息
		• ds-controller使用
		• Faiss-Server
		• GKS
		• Kubeflow
		  ◦ 基于Kubeflow的Training示例
		  ◦ 基于Kubeflow或Seldon的在线服务
		  ◦ Kubeflow Easyrec Pipeline示例
		  ◦ 转换自定义DAG为Pipeline
		• 使用教程：
		  ◦ 分布式Inference解决方案
		  ◦ 分布式XGBoost训练
		  ◦ 提交Spark任务到Hadoop集群

Dataflow	是EMR平台上提供的实时计算一站式解决方案，拥有分布式的、高吞吐量和高可扩展性的消息系统Kafka和基于Apache Flink官方产品Ververica提供的Flink商业内核两大组件，专注于解决实时计算端到端的各类问题、广泛应用于实时数据ETL和日志采集分析等场景，您也可以单独使用其中任一组件。	• Flink • Kafka

集群	描述	重要操作

ClickHouse	是一个面向联机分析处理（OLAP）的开源的面向列式存储的数据库管理系统。	• ClickHouse概述
		• 快速入门
		  ○ 创建集群
		  ○ 快速使用ClickHouse
		  ○ 访问模式
		• ClickHouse运维
		  ○ 日志配置说明
		  ○ 系统表说明
		  ○ 监控
		  ○ 配置项说明
		  ○ 访问权限控制
		  ○ 扩容ClickHouse集群
		• 数据导入
		  ○ 从Spark导入数据至ClickHouse
		  ○ 从Flink导入数据至ClickHouse
		  ○ 从HDFS导入数据至ClickHouse
		  ○ 从RDS导入数据至ClickHouse
		  ○ 从Kafka导入数据至ClickHouse

Druid	提供半托管式实时交互式分析服务，大数据查询毫秒级延迟，支持多种数据摄入方式，可以与EMR Hadoop、EMR Spark、OSS和RDS等服务搭配组合使用，构建灵活稳健的实时查询解决方案。	• 创建集群 • 登录集群 • 释放集群

Presto	是一种开源的交互式查询引擎，提供SQL on everything的能力。用于快速分析查询任何规模的数据，可以支持非关系数据源，例如，Hadoop分布式文件系统 (HDFS)、OSS、HBase、MongoDB；也可以支持关系数据源，例如MySQL、PostgreSQL、Microsoft SQL Server 和 Teradata；还可以支持数据湖文件，例如，Iceberg和Hudi。	• 概述
		• 常用连接器
		  ○ Hive连接器
		  ○ Kudu连接器
		  ○ MySQL连接器
		  ○ Iceberg连接器

集群	描述	重要操作

EMR Studio	是EMR平台上基于开源组件的大数据开发平台，提供一站式的端到端大数据开发体验。	• EMR Studio概述
		• 快速入门
		• Zeppelin
		  ◦ Spark
		  ◦ Flink
		  ◦ Hive
		  ◦ Presto
		  ◦ Shell
		  ◦ TPCH和TPCDS
		• JupyterHub
		  ◦ 管理JupyterHub
		  ◦ 使用Python3 Kernel运行EMR PySpark
		• Airflow
		  ◦ 基础使用
		  ■ Airflow常用配置说明
		  ■ 管理DAG
		  ■ 配置Airflow报警事件
		  ■ 代码示例
		  ◦ 最佳实践
		  ■ 定期调度Zeppelin中的作业
		  ■ 定期调度Jupyter中的作业
		  ■ 动态启动计算集群运行工作流调度
		  ◦ 常见问题

组件

Hadoop集群 | Zookeeper集群 | DataScience集群 | Druid集群 | Dataflow集群 | Presto集群

组件	文档链接
HDFS	概述
YARN	概述
Hive	概述
Spark	概述
Knox	概述

Tez	概述
Sqoop	概述
SmartData	概述
OpenLDAP	概述
Hudi	概述
Hue	概述
HBase	概述
Zookeeper	概述
Presto	概述
impala	概述
Zeppelin	概述
Flume	概述
Livy	概述
Ranger	概述
Phoenix	概述
ESS	概述
Alluxio	概述
Kudu	概述
Oozie	概述

2.集群规划

2.1. 选型配置说明

选择合适的集群是E-MapReduce产品使用的第一步。E-MapReduce配置选型不仅要考虑企业大数据使用场景、估算数据量、服务可靠性要求，还应该考虑企业预算。

大数据使用场景

E-MapReduce产品当前主要满足企业的以下大数据场景：

- 批处理场景
  该场景对磁盘吞吐和网络吞吐要求高，处理的数据量也大，但对数据处理的实时性要求不高，您可以选用MapReduce、Pig和Spark组件。该场景对内存要求不高，选型时您需要重点关注作业对CPU和内存的需求，以及Shuffle对网络的需求。
- Ad-Hoc查询

数据科学家或数据分析师使用即席查询工具检索数据。该场景对查询实时性、磁盘吞吐和网络吞吐要求高，您可以选用E-MapReduce的Impala和Presto组件。该场景对内存要求高，选型时需要考虑数据和并发查询的数量。

- 流式计算、高网络吞吐和计算密集型场景
  您可以选用E-MapReduce的Flink、Spark Streaming和Storm组件。
- 消息队列
  该场景对磁盘吞吐和网络吞吐要求高，并且内存消耗大，存储不依赖于HDFS。您可以选用E-MapReduce的Kafka集群。
- 数据冷备场景
  该场景对计算和磁盘吞吐要求不高，但要求冷备成本低，推荐使用JindoFS将阿里云OSS的归档型和深度归档型作为冷数据存储，以降低存储成本。

E-MapReduce节点

E-MapReduce节点有主实例（Master）、核心实例（Core）和计算实例（Task）三种实例类型。详情请参见实例类型。

E-MapReduce存储可以采用高效云盘、本地盘、SSD云盘和SSD本地盘。磁盘性能为SSD本地盘 > SSD云盘 > 本地盘 > 高效云盘。

E-MapReduce底层存储支持OSS（仅标准型OSS）和HDFS。相对于HDFS，OSS的数据可用性更高（99.99999999%），HDFS的数据可用性由云盘或本地盘存储的可靠性来保证。归档数据和深度归档数据需要解冻为标准型存储才能参与EMR引擎计算。

存储价格估算如下：

- 本地盘实例存储为0.04 元/GB/月
- OSS标准型存储为0.12 元/GB/月
- OSS归档型存储为0.033 元/GB/月
- OSS深度归档型存储为0.015 元/GB/月
- 高效云盘存储为0.35 元/GB/月
- SSD云盘存储为1.00 元/GB/月

E-MapReduce选型

- Master节点选型
  - Master节点主要部署Hadoop的Master进程。例如，NameNode和ResourceManager等。
  - 生产集群建议打开高可用HA，E-MapReduce的HDFS、YARN、Hive和HBase等组件均已实现HA。生产集群建议在创建集群的**硬件配置**步骤开启高可用。如果购买时未开启高可用，在后续使用过程中无法开启高可用功能。
  - Master节点主要用来存储HDFS元数据和组件Log文件，属于计算密集型，对磁盘IO要求不高。HDFS元数据存储在内存中，建议根据文件数量选择16 GB以上内存空间。
- Core节点选型
  - Core节点主要用来存储数据和执行计算，运行DataNode和Nodemanager。
  - HDFS数据量大于60 TB，建议采用本地盘实例（ECS.d1，ECS.d1NE），本地盘的磁盘容量为（CPU核数/2）*5.5TB*实例数量。
    例如，购买4台8核D1实例，磁盘容量为8/2*5.5*4 台=88 TB。因为HDFS采用3备份，所以本地盘实例最少购买3台，考虑到数据可靠性和磁盘损坏因素，建议最少购买4台。
  - HDFS数据量小于60 TB，可以考虑高效云盘和SSD云盘。
- Task节点选型

Task节点主要用来补充Core节点CPU和内存计算能力的不足，节点并不存储数据，也不运行DataNode。您可以根据CPU和内存需求来估算实例个数。

E-MapReduce生命周期

E-MapReduce支持弹性扩展，可以快速的扩容，灵活调整集群节点配置，详情请参见扩容，或者升降配ECS节点，详情请参见升降配。

可用区选择

为保证效率，您应该部署E-MapReduce与业务系统在同一地域的同一个可用区。详情请参见地域和可用区。

2.2. 实例类型

E-MapReduce集群由多个不同类型的实例节点组成，包括主实例节点（Master）、核心实例节点（Core）和计算实例节点（Task）。
不同实例节点上部署的服务进程不同，负责完成的任务也不同。例如：

- 主实例节点（Master）：部署Hadoop HDFS的NameNode服务、Hadoop YARN的ResourceManager服务。
- 核心实例节点（Core）：部署DataNode服务、Hadoop YARN的NodeManager服务。
- 计算实例节点（Task）：只进行计算，部署Hadoop YARN的NodeManager服务，不部署任何HDFS相关的服务。

创建集群时，您需要确定对应的三种实例类型的ECS规格，相同实例类型的ECS在同一个实例组内。创建集群完成后，您可以通过扩容来增加实例组内的机器数量（主实例组除外）。

> **说明** EMR-3.2.0及后续版本支持计算实例节点（Task）。

主实例节点（Master）

主实例节点是集群服务部署管控等组件的节点，例如，Hadoop YARN的 ResourceManager。

当您需要查看集群上服务的运行情况时，您可以通过软件的Web UI来查看。当您需要快速测试或者运行作业时，您可以登录主实例节点，然后通过命令行直接提交作业。登录主节点的具体步骤请参见登录集群。

核心实例节点（Core）

核心实例节点是被主实例节点管理的节点。核心实例节点上会运行Hadoop HDFS的Datanode服务，并保存所有的数据。同时，核心实例节点也会部署计算服务来执行计算任务。例如，Hadoop YARN的NodeManager服务。

为满足存储数据量或计算量扩展的需求，核心实例节点支持随时扩容，并且扩容过程中不会影响当前集群的正常运行。核心实例节点可以使用多种不同的存储介质来保存数据，详情请参见本地盘和块存储。

计算实例节点（Task）

计算实例节点是专门负责计算的实例节点，不会保存HDFS数据，也不会运行Hadoop HDFS的Datanode服务，是一个可选的实例类型。如果核心实例的计算能力充足，则可以不使用计算实例。当集群计算能力不足时，您可以随时通过计算实例节点快速给集群增加额外的计算能力，例如Hadoop的MapReduce任务和Spark Executors等。

计算实例节点可以随时新增和减少，并且不会影响现有集群的运行。

2.3. Gateway实例说明

创建Gateway集群时必须关联到一个已经存在的集群。Gateway集群可以作为一个独立的作业提交点，以便您更好的对关联集群进行操作。
Gateway集群通常是一个独立的集群，由多台相同配置的节点组成，集群上会部署Hadoop（HDFS+YARN）、Hive、Spark和Sqoop等客户端。

未创建Gateway集群时，Hadoop等集群的作业是在本集群的Master或Core节点上提交的，会占用本集群的资源。创建Gateway集群后，您可以通过Gateway集群来提交其关联的集群的作业，这样既不会占用关联集群的资源，又可以提高关联集群Master或Core节点的稳定性，尤其是Master节点。

每一个Gateway集群均支持独立的环境配置。例如，在多个部门共用一个集群的场景下，您可以为这个集群创建多个Gateway集群，以满足不同部门的业务需求。

2.4. ECS实例说明

本文介绍E-MapReduce（简称EMR）支持的ECS实例类型，以及各实例类型适用的场景。

EMR支持的ECS实例类型

- 通用型
  vCPU : Memory = 1 : 4。例如，8核32 GiB，使用云盘作为存储。
- 计算型
  vCPU : Memory = 1 : 2。例如，8核16 GiB，使用云盘作为存储，提供了更多的计算资源。
- 内存型
  vCPU : Memory = 1 : 8。例如， 8核64 GiB，使用云盘作为存储，提供了更多的内存资源。
- 大数据型
  使用本地SATA盘作存储数据，存储性价比高，是大数据量（TB级别的数据量）场景下的推荐机型。

  > **说明** Hadoop、Data Science、Dataflow和Druid类型的集群支持Core节点；Zookeeper和Kafka类型的集群不支持Core节点。

- 本地SSD型
  使用本地SSD盘，具有高随机IOPS（Input/Output Operations Per Second）和高吞吐能力。
- 共享型（入门级）
  共享CPU的机型，在大计算量的场景下，稳定性不够。入门级学习使用，不推荐企业客户使用。
- GPU
  使用GPU的异构机型，可以用来运行机器学习等场景。

实例类型适用场景

- Master主实例
  适合通用型或内存型实例，数据直接使用阿里云的云盘来保存，确保了数据的高可靠性。
- Core核心实例
  - 小数据量（TB级别以下）或者是使用OSS作为主要的数据存储时，推荐使用通用型、计算型或内存型。
  - 大数据量（10 TB或以上）情况下，推荐使用大数据机型，可以获得极高的性价比。

  > **注意** 当Core核心实例使用本地盘时，HDFS数据存储在本地盘，需要您自行保证数据的可靠性。

- Task计算实例
  用于补充集群的计算能力，可以使用除大数据型外的所有机型。

2.5. 存储说明

本文介绍E-MapReduce集群中数据存储相关的信息，包括磁盘角色、云盘与本地盘，以及OSS。

背景信息

关于存储的类型、性能和相关的限制信息，请参见块存储概述。

磁盘角色

在E-MapReduce集群中，实例节点上有系统盘和数据盘两种角色的磁盘。每块磁盘的配置、类型和容量都可以不同。

磁盘角色	描述
系统盘	系统盘用于安装操作系统。 E-MapReduce默认使用ESSD云盘作为集群的系统盘。系统盘默认是一块。
数据盘	数据盘用于保存数据。 Master实例默认挂载1块云盘作为数据盘，Core实例默认挂载4块云盘作为数据盘。

云盘与本地盘

E-MapReduce集群支持使用以下两种类型的磁盘来存储数据。

块存储类型	描述	使用场景
云盘	包括SSD云盘、高效云盘和ESSD云盘。 磁盘不直接挂载在本地的计算节点上，而是通过网络访问远端的一个存储节点。每一份数据在后端都有两个实时备份，一共三份数据。当一份数据损坏时（磁盘损坏，不是业务上的破坏），E-MapReduce会自动使用备份数据进行恢复。	当业务数据量处于TB级别以下时，推荐您使用云盘，云盘的IOPS和吞吐相比本地盘都会小些。 **说明** 在使用云盘时，如果吞吐量明显不足，则可以新建集群以使用本地盘。
本地盘	包括大数据型的SATA本地盘和本地SSD盘。 磁盘直接挂载在计算节点上，性能高于云盘。本地盘不能选择磁盘数量，只能使用默认配置好的数量，数据也没有后端的备份机制，需要上层的软件来保证数据可靠性。	当数据量处于TB级别以上时，推荐您使用本地盘，本地盘的数据可靠性由E-MapReduce来保证。

在E-MapReduce集群中，当实例节点释放时，所有云盘和本地盘都会清除数据，磁盘无法独立的保存下来并再次使用。Hadoop HDFS会使用所有的数据盘作为数据存储。 Hadoop YARN也会使用所有的数据盘作为计算的临时存储。

OSS

在E-MapReduce集群中，您可以将OSS作为HDFS使用。 E-MapReduce可以方便的读写OSS上的数据，所有使用HDFS的代码经过简单的修改即可以访问OSS的数据。例如：

- 读取HDFS中的数据。

```
sc.textfile("hdfs://user/path")
```

- 更改存储类型为OSS。

```
sc.textfile("oss://user/path")
```

- 对于MR或Hive作业，HDFS命令可以直接操作OSS数据，示例如下。

```
hadoop fs -ls oss://bucket/path
hadoop fs -cp hdfs://user/path  oss://bucket/path
```

此过程中，您不需要输入AccessKey和Endpoint，E-MapReduce会使用当前集群所有者的信息自动补全AccessKey和Endpoint。但OSS的IOPS不高，不适合用在IOPS要求高的场景，例如，流式计算Spark Streaming和HBase。

2.6. 本地盘机型概述

为了满足大数据场景下的存储需求，阿里云在云上推出了D1系列本地盘机型。

D1系列

D1系列使用本地盘而非云盘作为存储，解决了之前使用云盘产生多份冗余数据而导致的高成本问题。D1系列的数据传输不需要全部通过网络，这样不仅提高了磁盘的吞吐能力，还能发挥Hadoop的就近计算的优势。相较于云盘，本地盘机型提高了存储性能，降低了存储单价，成本与线下物理机相同。

本地盘机型存在一个问题，即数据可靠性问题。对于云盘，阿里云默认具有磁盘多备份策略，您完全感知不到磁盘的损坏，云盘可以自动保证数据的可靠性。对于本地盘，数据可靠性需要由上层的软件来保证，并且磁盘与节点故障也需要人工进行运维处理。

EMR+D1方案

针对本地盘机型（例如D1），E-MapReduce产品推出了一整套的自动化运维方案，帮助您方便可靠的使用本地盘机型。使您不需要关心整个运维的过程，同时还保证了数据高可靠和服务高可用。

自动化运维方案的主要点如下：

- 强制节点的高可靠分布
- 本地盘与节点的故障监控
- 数据迁移时机自动决策
- 自动的故障节点迁移与数据平衡
- 自动的HDFS数据检测
- 网络拓扑调优

通过整个后台的管控系统的自动化运维，E-MapReduce可以协助您更好的使用本地盘机型，实现高性价比的大数据系统。

> **说明** 如需使用D1机型搭建Hadoop集群，请提交工单。

2.7. 集群容灾能力

本文介绍E-MapReduce集群数据容灾和服务容灾能力。

数据容灾

在Hadoop分布式文件系统（HDFS）中，每一个文件的数据均是分块存储的，每一个数据块保存有多个副本（默认为3），并且尽量保证这些数据块副本分布在不同的机架之上。一般情况下，HDFS的副本系数是3，存放策略是将一个副本存放在本地机架节点上，一个副本存放在同一个机架的另一个节点上，最后一个副本放在不同机架的节点上。

HDFS会定期扫描数据副本，如果扫描到有数据副本丢失，则会快速复制这些数据以保证数据副本的数量。如果扫描到节点丢失，则节点上的所有数据也会快速复制恢复。在阿里云上，如果使用的是云盘技术，则每一个云盘在后台都会对应三个数据副本，当其中任一个出现问题时，副本数据都会自动进行切换并恢复，以保证数据的可靠性。

Hadoop HDFS是一个经历了长时间考验且具有高可靠性的数据存储系统，已实现了海量数据的高可靠性存储。同时基于云上的特性，您也可以再在OSS等服务上额外备份数据，以达到更高的数据可靠性。

服务容灾

Hadoop的核心组件都会进行HA部署，即有至少两个节点的服务互备，例如YARN、HDFS、Hive Server和Hive Meta。在任何一时刻，任一服务节点故障时，当前的服务节点都会自动进行切换，以保证服务不受影响。

2.8. EMR Hive元数据介绍与对比

本文为您介绍E-MapReduce（简称EMR）支持的元数据类型和各元数据类型的优势。

元数据类型介绍

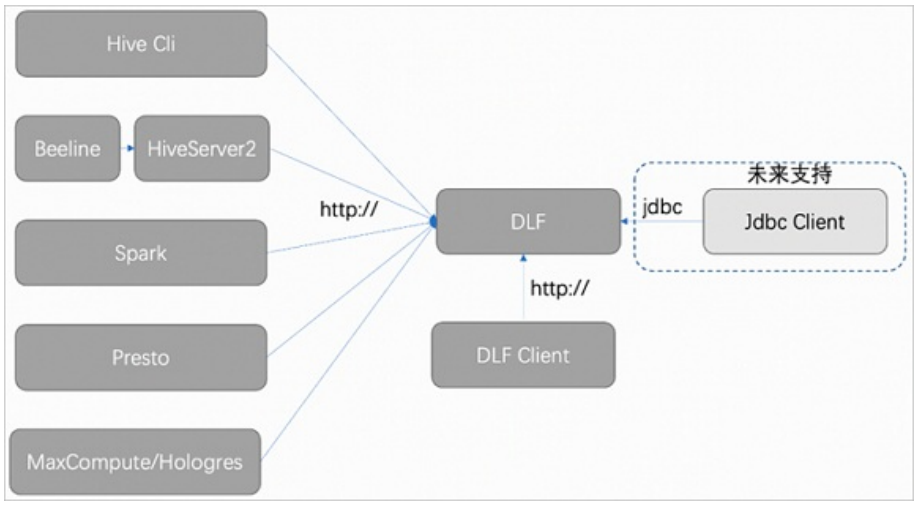
EMR Hive元数据支持DLF统一元数据、自建RDS和内置MySQL三种类型。

DLF统一元数据

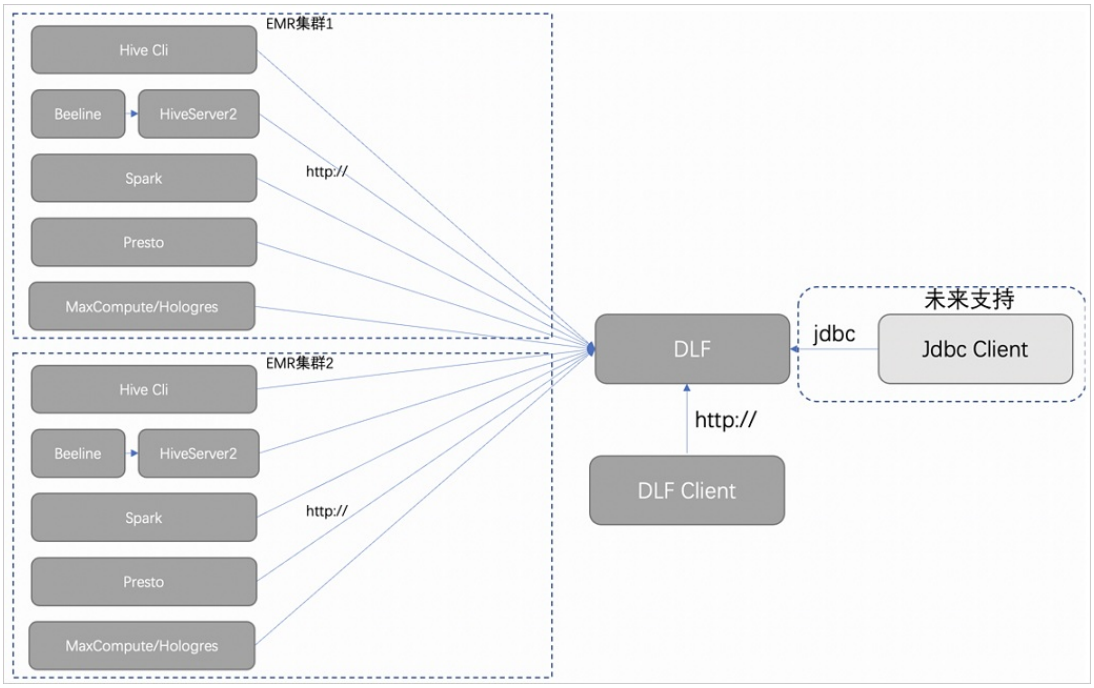
元数据存储在阿里云数据湖构建（Data Lake Formation，简称DLF）中。数据湖构建具有高可用、免运维和高性能等优点，兼容Hive Metastore，无缝对接EMR上开源计算引擎，并支持元数据多版本管理和Data Profile功能。另外，DLF还支持数据探索、湖管理和数据权限控制等功能，并与阿里云其他计算产品（例如MaxCompute、Databricks和Hologres等）无缝对接，可以扩展更丰富的计算场景，DLF详情请参见产品简介。

该元数据类型相比自建RDS和内置MySQL两种方式的最大区别是，无需在EMR集群上部署Hive Metastore，即元数据查询服务以及存储服务都托管到DLF产品上，免去运维成本，同时支持更多引擎（例如MaxCompute、Flink、DataBricks或Hologres等），进一步实现湖仓一体共享元数据，在多个集群上也能够实现元数据共享。DLF Client SDK提供了兼容Hive Metastore的接口，这样引擎基本不做任何改动就可以直接使用DLF Client SDK，进而访问DLF元数据。用户也可以直接使用DLF客户端访问DLF元数据。

DLF统一元数据在单集群部署架构图



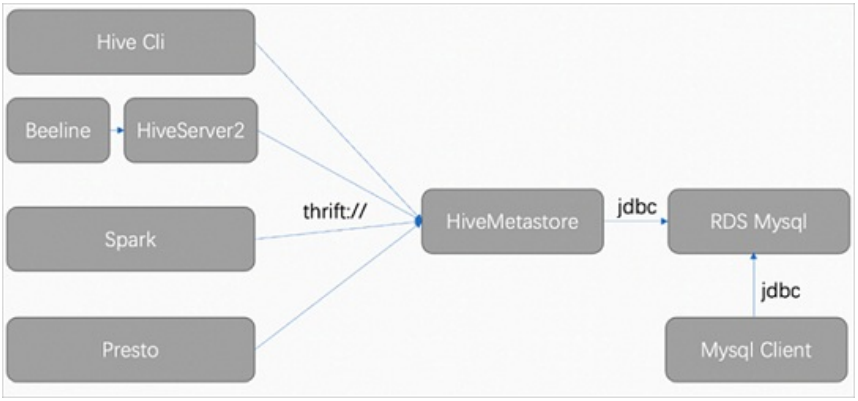
DLF统一元数据在多集群部署架构图



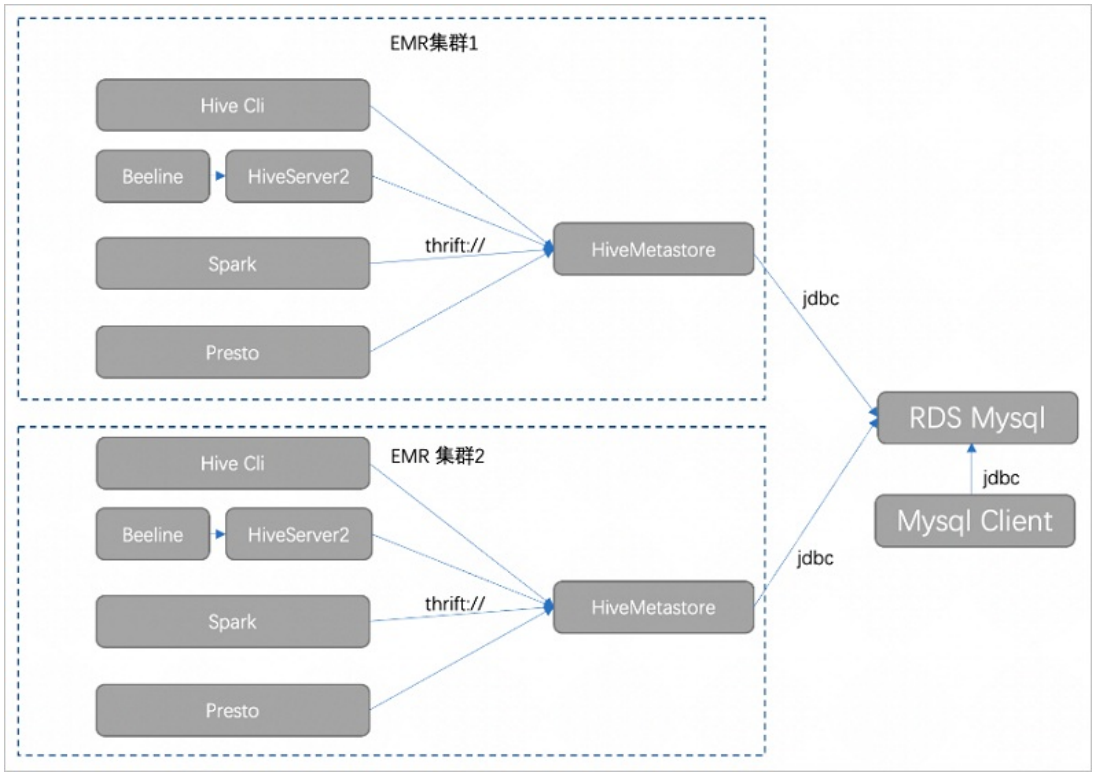
自建RDS

元数据存储在RDS中。自建RDS元数据类型和内置MySQL类型在架构上是一致的，区别只是存储由本地的MySQL，变成了RDS（云上的MySQL）。如自建RDS在多集群部署架构图所示，在多个集群环境中，RDS支持跨多个集群元数据共享，分别被不同集群中的Hive Metastore访问。

自建RDS在单集群部署架构图



自建RDS在多集群部署架构图

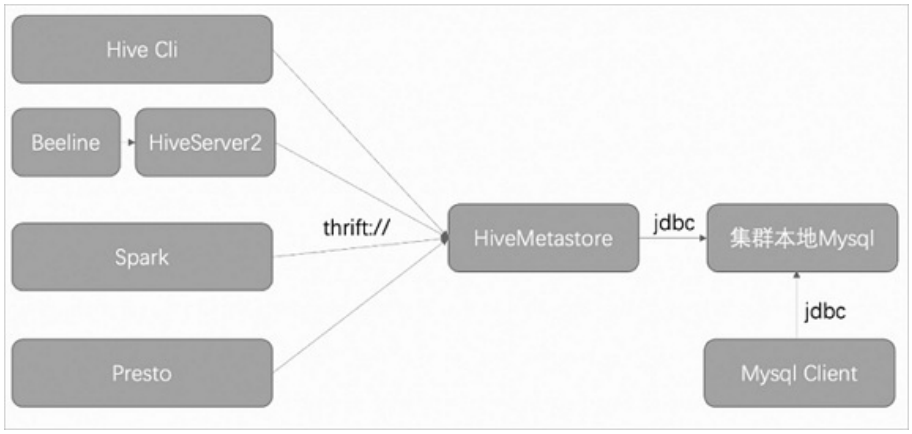


内置MySQL

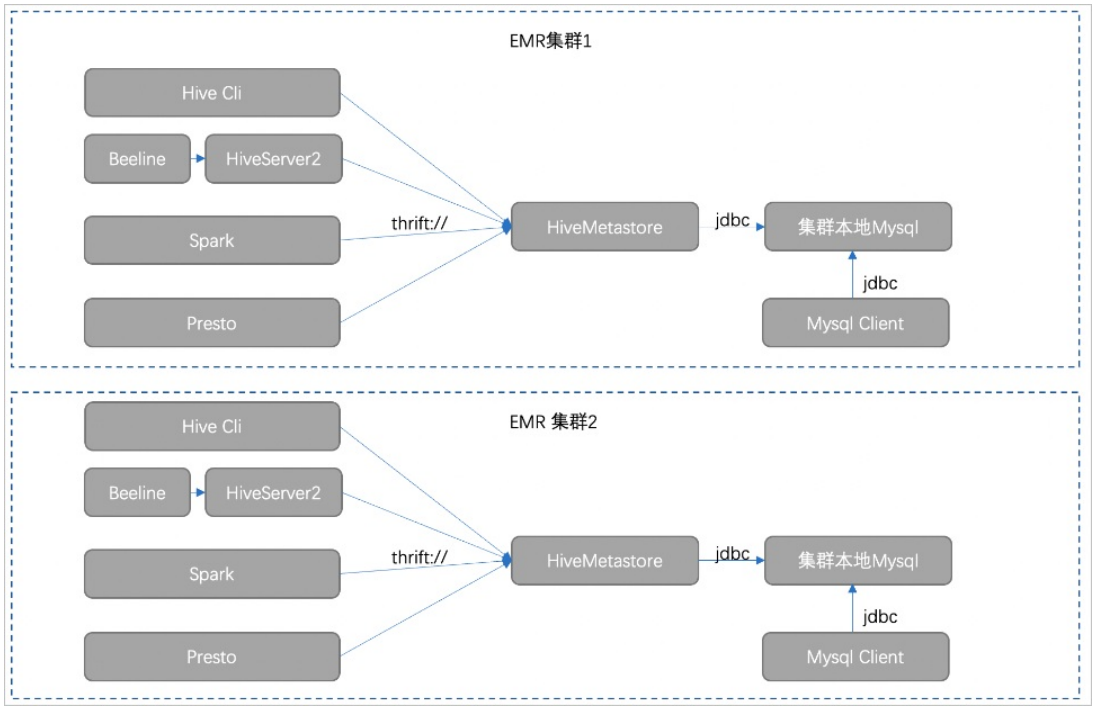
元数据存储在MySQL中，且MySQL Server实例部署在用户的EMR集群中（通常是Master节点），Hive、Spark或Presto等引擎访问元数据时，也不是直接访问MySQL，而是通过访问Hive Metastore间接访问MySQL。Hive Metastore提供元数据访问服务，引擎通过thrift协议访问Hive Metastore，Hive Metastore通过JDBC协议访问MySQL。

当然您也可以手动通过MySQL客户端直接连接MySQL Server，查看元数据。如内置MySQL在多集群部署架构图所示，由于每个集群都有一个MySQL，导致多个集群间的元数据不能共享。

内置MySQL在单集群部署架构图



内置MySQL在多集群部署架构图



元数据类型优势

内置MySQL和自建RDS的区别

自建RDS更直观的好处是元数据可以在多个集群间共享。

从可用性、可靠性和性能等方面对比，自建RDS要优于内置MySQL，详情请参见RDS与自建数据库对比优势。

DLF统一元数据和自建RDS的区别

对比项	DLF统一元数据	自建RDS
易用性	EMR集群开箱即用（需提前开通DLF产品）。	EMR集群开箱即用（需提前购买RDS）。
元数据管理	DLF产品提供了可视化的元数据检索、元数据管理、多版本管理、数据统计概况和生命周期管理等更丰富的能力。	无。
多引擎支持	• 支持Hive、Spark和Presto。 • 支持MaxCompute和Hologres。	• 支持Hive、Spark和Presto。 • 不支持MaxCompute和Hologres。
成本	免费，具体计费请参见数据湖构建的计费模式。	包月和按量计费。
运维成本	免运维，自动水平弹性扩容。	需要进行升级和扩容等运维操作。
高可用	主备等容灾措施。	主备等容灾措施。
性能	性能高，同时基于Hive Metastore进行了优化。	性能高，对Hive Metastore无优化。
更多能力	细粒度数据权限控制、湖存储分析和湖格式管理。	无。

非EMR集群访问DLF元数据

非EMR集群（本地测试环境或者其它云服务）访问DLF元数据，需要集成DLF Client SDK，具体操作请参见阿里云数据湖构建（DLF）。

说明 访问DLF和访问MySQL一样，需要提供访问地址、用户名和密码。

- DLF中的访问地址称为Endpoint，不同region使用不同的Endpoint。
- 用户名和密码为DLF产品的AccessKey ID和AccessKey Secret。
- 您还需要通过修改Hive参数的方式，切换Hive MetaStore的存储方式。
  在Hive服务页面，修改参数hive.imetastoreclient.factory.class的值为com.aliyun.datalake.metastore.hive2.DlfMetaStoreClientFactory。修改配置项的具体操作请参见修改配置项。

3.集群配置

3.1. 创建集群

本文介绍创建E-MapReduce（简称EMR）集群的详细操作步骤和相关配置。

前提条件

已完成RAM授权，操作步骤请参见角色授权。

操作步骤

1. 进入创建集群页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
      - 地域：创建的集群将会在对应的地域内，一旦创建不能修改。
      - 资源组：默认显示账号全部资源。
   iii. 单击**创建集群**，进行创建。
2. 配置集群信息。

   创建集群时，您需要对集群进行软件配置、硬件配置和基础配置。

   > 注意 集群创建完成后，除了集群名称以外，其他配置均无法修改，所以在创建时请仔细确认各项配置。

   i. 软件配置。

配置项	说明

配置项	说明

集群类型	当前支持的集群类型如下：
	■ **Hadoop**：
	  ■ 提供半托管的Hadoop、Hive和Spark离线大规模分布式数据存储和计算。
	  ■ 提供SparkStreaming、Flink和Storm流式数据计算。
	  ■ 提供Presto和Impala交互式查询。
	  ■ 提供Oozie和Pig等Hadoop生态圈的组件。
	■ **Zookeeper**：提供独立的分布式一致性锁服务，适用于大规模的Hadoop集群、HBase集群和Kafka集群。
	■ **Data Science**：主要面向大数据+AI场景，提供Hive和Spark离线大数据ETL和TensorFlow模型训练，您可以选择CPU+GPU的异构计算框架，通过英伟达GPU对部分深度学习算法进行高性能计算。
	■ **Druid**：提供半托管式实时交互式分析服务，大数据查询毫秒级延迟，支持多种数据摄入方式，可以与EMR Hadoop、EMR Spark、OSS和RDS等服务搭配组合使用，构建灵活稳健的实时查询解决方案。
	■ **ClickHouse**：是一个面向联机分析处理（OLAP）的开源的面向列式存储的数据库管理系统。
	■ **Dataflow**：是EMR平台上提供的实时计算一站式解决方案，拥有分布式的、高吞吐量和高可扩展性的消息系统Kafka和基于Apache Flink官方产品Ververica提供的Flink商业内核两大组件，专注于解决实时计算端到端的各类问题、广泛应用于实时数据ETL和日志采集分析等场景，您也可以单独使用其中任一组件。
	■ **Presto**：是一种开源的交互式查询引擎，提供SQL on everything的能力。用于快速分析查询任何规模的数据，可以支持非关系数据源。
	■ **EMR Studio**：是EMR平台上基于开源组件的大数据开发平台，提供一站式的端到端大数据开发体验。更多信息，请参见EMR Studio概述。

云原生选项	默认on ECS。
产品版本	默认最新的软件版本。
必选服务	默认的服务组件，后期可以在管理页面中启停服务。
可选服务	根据您的实际需求选择其他的一些组件，被选中的组件会默认启动相关的服务进程。 **说明** 组件越多，对机器的配置要求也越高，所以在下面的步骤中您需要根据实际的组件数量进行机器选型，否则可能没有足够的资源运行这些服务。
高级设置	■ **Kerberos集群模式**：是否开启集群的Kerberos认证功能。默认不开启。通常个人用户集群无需该功能。 ■ **软件自定义配置**：可指定JSON文件对集群中的基础软件（例如Hadoop、Spark和Hive等）进行配置，详细使用方法请参见软件配置。默认不开启。

ii. 硬件配置。

区域	配置项	说明
付费类型	**付费类型**	默认包年包月。当前支持的付费类型如下： ■ **按量付费**：一种后付费模式，即先使用再付费。按量付费是根据实际使用的小时数来支付费用，每小时计费一次，适合短期的测试任务或是灵活的动态任务。 ■ **包年包月**：一种预付费模式，即先付费再使用。 **说明** ■ 建议测试场景下使用**按量付费**，测试正常后再新建一个**包年包月**的生产集群正式使用。 ■ 包年包月实例还需选择付费时长和是否开启自动续费。默认续费时长为1个月，且未开启自动续费。开启自动续费后，实例到期前7天会执行自动续费操作，续费时长为1个月，详情请参见续费流程。
网络配置	**可用区**	可用区为在同一地域下的不同物理区域，可用区之间内网互通。通常使用默认的可用区即可。
	网络类型	默认专有网络。
	VPC	选择在该地域的VPC。如果没有可用的VPC，单击**创建VPC/子网（交换机）**前往新建。
	交换机	选择在对应VPC下可用区的交换机，如果在这个可用区没有可用的交换机，则需要新创建一个。
	安全组名称	默认选择已有的安全组。安全组详情请参见安全组概述。 您也可以单击**新建安全组**，然后直接输入安全组名称来新建一个安全组。 注意 禁止使用ECS上创建的企业安全组。
高可用	**高可用**	默认不开启。打开**高可用**开关，Hadoop集群会有两个或三个Master节点来支持ResourceManager和NameNode的高可用。 HBase集群原本就支持高可用，只是另一个节点用其中一个Core节点来充当，如果打开高可用，会独立使用一个Master节点来支持高可用，更加的安全可靠。

区域	配置项	说明
实例	选型配置	■ **Master实例**：主要负责ResourceManager和NameNode等控制进程的部署。 您可以根据需要选择实例规格，详情请参见实例规格族。 ■ **系统盘配置**：根据需要选择SSD云盘、ESSD云盘或者高效云盘。 ■ **系统盘大小**：根据需要调整磁盘容量，推荐至少120 GB。取值范围为40 ~ 2048 GB。 ■ **数据盘配置**：根据需要选择SSD云盘、ESSD云盘或者高效云盘。 ■ **数据盘大小**：根据需要调整磁盘容量，推荐至少80 GB。取值范围为40 ~ 32768 GB。 ■ **Master数量**：默认1台。如果开启高可用默认2或者3台。 ■ **Core实例**：主要负责集群所有数据的存储，创建集群完成后也支持按需进行扩容。 ■ **系统盘配置**：根据需要选择SSD云盘、ESSD云盘或者高效云盘。 ■ **系统盘大小**：根据需要调整磁盘容量，推荐至少120 GB。 ■ **数据盘配置**：根据需要选择SSD云盘、ESSD云盘或者高效云盘。 ■ **数据盘大小**：根据需要调整磁盘容量，推荐至少80 GB。 ■ **Core数量**：默认2台，根据需要调整。 ■ **Task实例**：不保存数据，调整集群的计算力使用。默认不开启，需要时再追加。

iii. 基础配置。

区域	配置项	说明
	集群名称	集群的名字，长度限制为1~64个字符，仅可使用中文、字母、数字、短划线（-）和下划线（_）。

区域	配置项	说明
基础信息	元数据选择	■ **DLF统一元数据**（推荐）：表示元数据存储在数据湖中。 数据湖构建（Data Lake Formation，DLF）的元数据管理可以为您提供全托管、免运维、高可用、高性能的统一元数据服务，并且兼容Hive多版本，可以方便的进行HMS间元数据迁移。阿里云数据湖构建的详细信息，请参见产品简介。 **说明** 如果您希望将Hive Metastore的元数据迁移到数据湖构建（DLF）中，详情请参见元数据迁移。 ■ **使用自建RDS**：表示使用自建的阿里云RDS作为元数据库，更多信息请参见配置独立RDS MySQL。 ■ **集群内置MySQL**（不推荐）：表示元数据存储在集群本地环境的MySQL数据库中。 **说明** 该方式仅限在测试场景下使用，因为本地MySQL数据库部署在EMR集群单节点中，不能保证服务高可用，存在稳定性风险。生产场景建议选择**DLF统一元数据**或**使用自建RDS**方式。
	挂载公网	集群是否挂载弹性公网IP地址，默认不开启。 **说明** 创建后如果您需要使用公网IP地址访问，请在ECS上申请开通公网IP地址，详情请参见弹性公网IP中的申请EIP的内容。
	密钥对	关于密钥对的使用详情，请参见SSH密钥对。
	密码	设置Master节点的登录密码，密码规则：8~30个字符，且必须同时包含大写字母、小写字母、数字和特殊字符。 特殊字符包括：感叹号（!）、at（@）、井号（#）、美元符号（$）、百分号（%）、乘方（^）、and（&）和星号（*）。
高级设置	添加用户	添加访问开源大数据软件Web UI的账号。
	权限设置	通过RAM角色为在集群上运行的应用程序提供调用其他阿里云服务所需的必要权限，无需调整，使用默认即可。 ■ **服务角色**：用户将权限授予EMR服务，允许EMR代表用户调用其他阿里云的服务，例如ECS和OSS。 ■ **ECS应用角色**：当用户的程序在EMR计算节点上运行时，可不填写阿里云AccessKey来访问相关的云服务（例如OSS），EMR会自动申请一个临时AccessKey来授权本次访问。**ECS应用角色**用于控制这个AccessKey的权限。

区域	配置项	说明
	数据盘加密	默认不开启。 打开**加密开关**，即启动对集群节点ECS中所有属性为云盘的数据盘进行加密的功能。默认使用服务密钥为用户的数据进行加密，也支持使用用户自选密钥为用户的数据进行加密。 注意 不支持加密本地盘。
	引导操作	可选配置，您可以在集群启动Hadoop前执行您自定义的脚本，详情请参见引导操作。
	标签	可选配置，您可以在创建集群时绑定标签，也可以在集群创建完成后，在集群详情页绑定标签，详情请参见设置标签。
	资源组	可选配置。详情请参见使用资源组。

说明 页面右边会显示您所创建集群的配置清单以及集群费用。根据不同的付费类型，展示不同的价格信息。

3. 当所有的信息确认正确有效后，选中服务条款，单击**创建**。

注意

- 按量付费集群：立刻开始创建。
  集群创建完成后，集群的状态变为**空闲**。
- 包年包月集群：先生成订单，在支付完成订单以后集群才会开始创建。

3.2. 创建Gateway集群

您可以通过Gateway集群实现负载均衡和安全隔离，也可以通过Gateway集群向E-MapReduce集群提交作业。本文介绍如何创建Gateway集群。

前提条件

已经在E-MapReduce中创建了Hadoop或Kafka类型的集群，详情请参见创建集群。

使用限制

在E-MapReduce中，Gateway集群仅支持关联Hadoop或Kafka类型的集群。

操作步骤

1. 进入**集群管理**页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
2. 在**集群管理**页面，单击右上角的**创建Gateway**。
3. 在**创建Gateway**页面，配置各参数。

区域	配置项	说明
基础信息	**集群名称**	Gateway集群的名称，长度限制为1~64个字符，只允许包含中文、字母、数字、短划线（-）、下划线（_）。
	挂载公网	Gateway是否挂载弹性公网IP地址。
	密码和密钥对	○ **登录密码**：在文本框中输入登录Gateway的密码。长度限制为8~30个字符，且必须同时包含大写字母、小写字母、数字和特殊字符。 支持输入以下字符： !@#$%^&* ○ **密钥对**：在列表中选择登录Gateway的密钥对名称。如果还未创建过密钥对，则您可以单击后面的**新建密钥对**，进入ECS管理控制台进行创建。 请妥善保管好密钥对所对应的私钥文件（*.pem*文件）。 Gateway创建成功后，该密钥对的公钥部分会自动绑定到Gateway所在的云服务器ECS上，当通过SSH登录Gateway时，您需要输入私钥文件中的私钥。
	付费类型	○ **包年包月**：一次性支付一段时间的费用，价格相对来说会比较便宜，特别是包三年时，折扣会比较大。 ○ **按量付费**：根据实际使用的小时数来支付费用，每个小时计一次费用。
集群配置	**关联集群**	Gateway集群需要关联的集群，即创建的Gateway可以提交作业的集群。
	可用区	关联集群所在的可用区（Zone）。
	网络类型	关联集群的网络类型。
	VPC	关联集群所在地域下的VPC。
	交换机	选择在对应的VPC下对应可用区的交换机。
	安全组名称	关联集群所属的安全组。

区域	配置项	说明
实例	**Gateway实例**	该地域内可选择的ECS实例规格，详细说明请参见实例规格族。 ○ **系统盘配置**：Gateway节点使用的系统盘类型。系统盘有高效云盘、ESSD云盘和SSD云盘三种，根据不同机型和不同的Region，系统盘显示类型会有不同。系统盘默认随着集群的释放而释放。 ○ **系统盘大小**：每块最小为40 GB，最大为500 GB。默认值为300 GB。 ○ **数据盘配置**：Gateway节点使用的数据盘类型。数据盘有高效云盘、ESSD云盘和SSD云盘三种，根据不同机型和不同的Region，数据盘显示类型会有不同。数据盘默认随着集群的释放而释放。 ○ **数据盘大小**：每块最小为200 GB，最大为4000 GB。默认值为300 GB。 ○ **数量**：数据盘的数量，最小设置为1台，最大设置为10台。
高级设置	**权限设置**	通过RAM角色为在集群上运行的应用程序提供调用其他阿里云服务所需的必要权限，无需调整，使用默认即可。 ○ **服务角色**：用户将权限授予EMR服务，允许EMR代表用户调用其他阿里云的服务，例如ECS和OSS。 ○ **ECS应用角色**：当用户的程序在EMR计算节点上运行时，可不填写阿里云AccessKey来访问相关的云服务（例如OSS），EMR会自动申请一个临时AccessKey来授权本次访问。**ECS应用角色**用于控制这个AccessKey的权限。
	引导操作	可选配置，您可以在集群启动前执行您自定义的脚本，详情请参见引导操作。
	数据盘加密	默认不开启。 打开**加密开关**，即启动对集群节点ECS中所有属性为云盘的数据盘进行加密的功能。默认使用服务密钥为用户的数据进行加密，也支持使用用户自选密钥为用户的数据进行加密。 注意 不支持加密本地盘。

4. 完成上述参数配置后，选中**E-MapReduce服务条款**，单击**创建**。
   创建成功后，集群的**状态**变为**空闲**。

3.3. 管理权限

3.3.1. 角色授权

阿里云E-MapReduce服务（例如Hadoop和Spark），在运行时需要有访问其他阿里云资源和执行操作的权限。每个E-MapReduce集群必须有服务角色以及ECS应用角色。本文为您介绍EMR角色授权的流程及其关联的角色。

背景信息

阿里云E-MapReduce为确定权限的角色提供默认角色和默认系统策略。系统策略由阿里云创建和维护，因此如果服务要求发生变化，策略会自动更新。

首次使用E-MapReduce服务时，您需要使用阿里云账号为E-MapReduce服务授权名为AliyunEMRDefaultRole、AliyunECSInstanceForEMRRole或AliyunEmrEcsDefaultRole的服务角色。授权成功后，您可以在RAM控制台上查看角色，并为角色挂载策略。角色详细信息，请参见RAM角色概览。

> 注意
>
> - E-MapReduce不同版本需要授权的服务角色名称不同：
>   - EMR-3.32.0及之前版本和EMR-4.5.0及之前版本：AliyunEmrEcsDefaultRole
>   - EMR-3.32.0之后版本和EMR-4.5.0之后版本：AliyunECSInstanceForEMRRole
> - 首次使用E-MapReduce服务时，必须使用阿里云账号完成默认角色授权，否则RAM用户和阿里云账号不能使用E-MapReduce。
> - 如果删除服务角色，请确保使用该角色的资源已经释放，否则会影响资源的正常使用。
> - 如果只授权了部分角色，E-MapReduce控制台会提醒授权，只有在全部角色授权完成之后才可以创建集群。

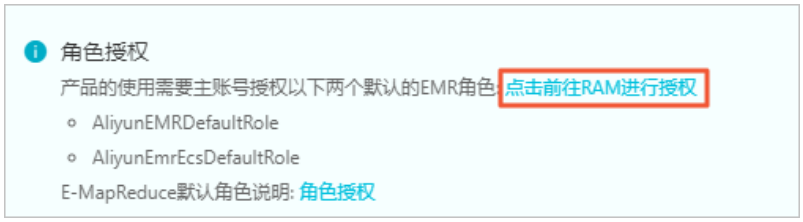
角色授权流程

本流程以EMR-3.30版本为例。

1. 在阿里云E-MapReduce控制台，单击**点击前往RAM进行授权**。

   > 说明 首次使用E-MapReduce服务时需要授权，授权成功后，再次使用无需重复授权。

   如果未正确地给E-MapReduce的服务账号授予默认角色，则在创建集群或创建按需执行计划时，会弹出如下提示。

2. 单击**同意授权**，将默认角色AliyunEMRDefaultRole和AliyunEmrEcsDefaultRole授予给E-MapReduce服务。

3. 完成以上授权后，您需要刷新E-MapReduce控制台，即可进行相关操作。

   如果您想查看AliyunEMRDefaultRole和AliyunEmrEcsDefaultRole相关的详细策略信息，您可以登录RAM的控制台查看。

服务角色

下表列出了与E-MapReduce关联的RAM角色。

属性	默认角色	描述	系统策略
EMR服务角色	AliyunEMRDefaultRole	允许E-MapReduce服务在配置资源和执行服务级别操作时调用其他阿里云服务。所有集群都需要该角色，且不能更改。 详情请参见EMR服务角色。	AliyunEMRRolePolicy
ECS应用角色（EMR 3.32及之前版本和EMR 4.5及之前版本）	AliyunEmrEcsDefaultRole	集群实例上运行的应用程序进程在调用其他阿里云服务时将使用该角色。在创建集群时既可以使用该服务角色，也可以使用自定义的角色。 AliyunEmrEcsDefaultRole角色，详情请参见ECS应用角色（EMR 3.32及之前版本和EMR 4.5及之前版本）。	AliyunEMRECSRolePolicy
ECS应用角色（EMR 3.32之后版本和EMR 4.5之后版本）	AliyunECSInstanceForEMRRole	集群实例上运行的应用程序进程在调用其他阿里云服务时将使用该角色。在创建集群时既可以使用该服务角色，也可以使用自定义的角色。 AliyunECSInstanceForEMRRole角色，详情请参见ECS应用角色（EMR 3.32之后版本和EMR 4.5之后版本）。	AliyunECSInstanceForEMRRolePolicy
ECS应用角色（EMR Studio默认使用）	AliyunECSInstanceForEMRStudioRole	EMR Studio使用此角色来访问您在其他云产品中的资源。 当您的账号未授权该角色，首次创建EMR Studio集群时会弹出授权该角色窗口，请使用阿里云账号授权该角色。	AliyunECSInstanceForEMRStudioRolePolicy

3.3.2. EMR服务角色

EMR服务角色允许E-MapReduce服务在配置资源或执行服务级别操作时调用其他阿里云服务。例如，服务角色用于在EMR集群启动时创建ECS实例。本文为您介绍EMR服务角色AliyunEMRDefaultRole及其权限策略。

注意事项

为了避免影响EMR服务稳定性，请注意：

- EMR服务角色名称无法修改。
- 不要在RAM访问控制台上删除或修改EMR服务角色的系统策略。

权限内容

AliyunEMRDefaultRole是默认的EMR服务角色，该服务角色包含**AliyunEMRRolePolicy**系统权限策略。其中**AliyunEMRRolePolicy**包含的权限内容如下：

- ECS相关权限

名称（Action）	说明
ecs:CreateInstance	创建ECS实例。
ecs:RenewInstance	为ECS实例续费。

名称（Action）	说明
ecs:DescribeRegions	查询ECS地域信息。
ecs:DescribeZones	查询Zone信息。
ecs:DescribeImages	查询镜像信息。
ecs:CreateSecurityGroup	创建安全组。
ecs:AllocatePublicIpAddress	分配公网IP地址。
ecs:DeleteInstance	删除ECS实例。
ecs:StartInstance	启动ECS实例。
ecs:StopInstance	停止ECS实例。
ecs:DescribeInstances	查询ECS实例。
ecs:DescribeDisks	查询ECS相关磁盘信息。
ecs:AuthorizeSecurityGroup	设置安全组入规则。
ecs:AuthorizeSecurityGroupEgress	设置安全组出规则。
ecs:DescribeSecurityGroupAttribute	查询安全组详情。
ecs:DescribeSecurityGroups	查询安全组列表信息。

- OSS相关权限

名称（Action）	说明
oss:PutObject	上传文件或文件夹对象。
oss:GetObject	获取文件或文件夹对象。
oss:ListObjects	查询文件列表信息。

3.3.3. ECS应用角色（EMR 3.32及之前版本和EMR 4.5及之前版本）

E-MapReduce环境提供了MetaService服务，MetaService服务是一种特殊的ECS应用角色。EMR 3.32及之前版本和EMR 4.5及之前版本，创建时会自动绑定该角色。在EMR集群之上运行的应用程序通过该角色来获得与其他云服务交互的权限，实现以免AccessKey的方式访问阿里云资源，避免了在配置文件中暴露AccessKey的风险。

前提条件

已授权该角色，详情请参见角色授权。

背景信息

当前MetaService服务仅支持免AccessKey访问OSS、LogService和MNS数据。

权限内容

默认服务角色**AliyunEmrEcsDefaultRole**包含系统权限策略为**AliyunEmrECSRolePolicy**，OSS相关权限内容如下。

权限名称（Action）	权限说明
oss:PutObject	上传文件或文件夹对象。
oss:GetObject	获取文件或文件夹对象。
oss:ListObjects	查询文件列表信息。
oss:DeleteObject	删除某个文件。
oss:AbortMultipartUpload	终止MultipartUpload事件。

注意 请谨慎编辑和删除默认角色AliyunEmrEcsDefaultRole，否则会造成集群创建失败或作业运行失败。

支持MetaService的数据源

目前在E-MapReduce上支持MetaService的产品有OSS、LogService和MNS。您可以在E-MapReduce集群上使用E-MapReduce SDK接口免AccessKey读写上述数据源。

MetaService默认只有OSS的读写权限，如果您希望MetaService支持LogService和MNS，请前往RAM控制台为AliyunEmrEcsDefaultRole应用角色增加LogService和MNS的读写权限，RAM控制台链接为RAM控制台。

RAM角色授权，详情请参见为RAM角色授权。

使用MetaService

基于MetaService服务，您可通过E-MapReduce作业免AccessKey访问阿里云资源（OSS、LogService和MNS），其优势如下：

- 降低AccessKey泄漏的风险。基于RAM，您可按最小够用原则给角色授权，做到权限最小化，这样可以将安全风险降到最低。
- 提高用户体验。尤其在交互式访问OSS资源时，可避免让您输入一长串的OSS路径。
- EMR自带服务
  EMR自带服务中运行的作业均可以自动基于MetaSerivce服务免明文AccessKey访问阿里云资源（OSS、LogService和MNS）
  以下是使用MetaService（新）和不使用MetaService（旧）的对比示例：
  - 通过Hadoop命令行查看OSS数据
    - 旧方式

```
hadoop fs -ls oss://ZaH******As1s:Ba23N**************sdaBj2@bucket.oss-cn-hangzhou-internal.aliyuncs.com/a/b/c
```

    - 新方式

```
hadoop fs -ls oss://bucket/a/b/c
```

- 通过Hive建表
  - 旧方式

```
CREATE EXTERNAL TABLE test_table(id INT, name string)
        ROW FORMAT DELIMITED
        FIELDS TERMINATED BY '/t'
        LOCATION 'oss://ZaH******As1s:Ba23N**************sdaBj2@bucket.oss-cn-hangzhou
-internal.aliyuncs.com/a/b/c';
```

  - 新方式

```
CREATE EXTERNAL TABLE test_table(id INT, name string)
        ROW FORMAT DELIMITED
        FIELDS TERMINATED BY '/t'
        LOCATION 'oss://bucket/a/b/c';
```

- 使用Spark查看OSS数据
  - 旧方式

```
val data = sc.textFile("oss://ZaH******As1s:Ba23N**************sdaBj2@bucket.oss-cn-ha
ngzhou-internal.aliyuncs.com/a/b/c")
```

  - 新方式

```
val data = sc.textFile("oss://bucket/a/b/c")
```

- 自行部署服务
  MetaService是一个HTTP服务，您可以直接访问这个HTTP服务的URL获取STS临时凭证，然后在自行搭建的系统中使用该STS临时凭证免AccessKey访问阿里云资源。

  > 注意 STS临时凭证失效前半小时会生成新的STS临时凭证，在这半小时内，新旧STS临时凭证均可使用。

  例如，通过**curl http://localhost:10011/cluster-region**，即可获得当前集群所在Region。
  当前MetaService支持以下几类信息：
  - Region：/cluster-region
  - 角色名：/cluster-role-name
  - AccessKeyID：/role-access-key-id
  - AccessKeySecret：/role-access-key-secret
  - SecurityToken：/role-security-token
  - 网络类型：/cluster-network-type

3.3.4. ECS应用角色（EMR 3.32之后版本和EMR 4.5之后版本）

EMR 3.32之后版本和EMR 4.5之后版本，将Metaservice服务替换为ECS应用角色，在EMR集群创建和扩容时自动分配给EMR集群中的每个ECS实例。在EMR集群之上运行的应用程序通过该角色来获得与其他云服务交互的权限，实现以免AccessKey的方式访问阿里云资源，避免了在配置文件中暴露AccessKey的风险。

前提条件

已授权该角色，详情请参见角色授权。

权限内容

默认角色AliyunECSInstanceForEMRRole包含系统权限策略为AliyunECSInstanceForEMRRolePolicy，OSS相关权限内容如下。

权限名称（Action）	权限说明
oss:PutObject	上传文件或文件夹对象。
oss:GetObject	获取文件或文件夹对象。
oss:ListObjects	查询文件列表信息。
oss:DeleteObject	删除某个文件。
oss:AbortMultipartUpload	终止MultipartUpload事件。

> **注意** 请谨慎编辑和删除默认角色AliyunEmrEcsDefaultRole，否则会造成集群创建失败或作业运行失败。

使用ECS应用角色获取临时凭证

基于STS临时凭证可以获取本账号下其他云资源的访问权限，详情请参见使用实例RAM角色访问其他云产品。

3.3.5. 使用自定义ECS应用角色访问同账号云资源

本文介绍在E-MapReduce控制台上，通过创建集群时在基础配置页面的高级设置区域设置ECS应用角色，实现以免密的方式访问同账号下的其它资源。例如，对象存储OSS和日志服务SLS。

背景信息

您在创建集群时可以使用自定义的角色，通过给该角色不同的权限策略，以限制集群访问外部资源的权限。例如，您可以进行如下操作：

- 指定集群只能访问指定OSS的数据目录。
- 指定集群访问指定的外部资源。

前提条件

- 已创建EMR集群，详情请参见创建集群。
- 已完成角色授权，详情请参见角色授权。
- 已在OSS管理控制台，创建与EMR集群同一地域下的存储空间，详情请参见创建存储空间。

操作流程

1. 步骤一：新建权限策略
2. 步骤二：创建RAM角色
3. 步骤三：创建集群并访问外部资源

步骤一：新建权限策略

1. 进入新建自定义权限策略页面。
    i. 使用云账号登录RAM控制台。

ii. 在RAM访问控制页面，选择**权限管理 > 权限策略**。

iii. 在**权限策略**页面，单击**创建权限策略**。

2. 在**新建自定义权限策略**页面，配置以下信息。

参数	描述
策略名称	本示例为test-emr。
配置模式	选择**脚本配置**。
策略内容	添加如下策略。 `{ "Version": "1", "Statement": [ { "Action": [ "oss:GetObject", "oss:ListObjects" ], "Resource": [ "acs:oss:*:*:emr-logs2", "acs:oss:*:*:emr-logs2/*" ], "Effect": "Allow" } ] }` **说明** 策略中涉及的元素如下所示： ○ Action：是指对具体资源的操作。本示例是OSS的读取和查询目录的权限。 ○ Resource：是指被授权的具体对象。本示例是访问*test-emr*的OSS Bucket及其中的内容。 更多权限策略的基本元素，请参见权限策略基本元素。

3. 单击**确定**。

步骤二：创建RAM角色

1. 在RAM访问控制页面，选择**身份管理 > 角色**。
2. 在**角色**页面，单击**创建角色**。
3. 创建RAM角色。

i. 单击**阿里云服务**。

ii. 单击**下一步**。

iii. 在**配置角色**面板，配置以下信息。

参数	描述
角色名称	本示例为test-emr。
选择授信服务	选择**云服务器**。

iv. 单击**完成**。

4. （可选）修改授信服务。

> **注意** 如果您创建的集群是EMR 3.32之后版本、EMR 4.5之后版本或EMR 5.x及之后版本，则无需执行本步骤。

i. 在**角色**页面，单击刚创建的**角色名称**。

ii. 单击**信任策略管理**页签。

iii. 单击**修改信任策略**。

iv. 修改 `ecs.aliyuncs.com` 为 `emr.aliyuncs.com` 。

```
{
  "Statement": [
    {
      "Action": "sts:AssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": [
          "ecs.aliyuncs.com"
        ]
      }
    }
  ],
  "Version": "1"
}
```

v. 单击**确定**。

5. 添加相应权限。

i. 在**角色**页面，单击刚创建**角色名称**的操作列的**添加权限**。

ii. 在**添加权限**页面，单击**自定义策略**，添加新建的权限策略。

iii. 单击**确定**。

iv. 单击**完成**。

步骤三：创建集群并访问外部资源

1. 登录阿里云E-MapReduce控制台。
2. 在顶部菜单栏处，根据实际情况选择地域和资源组。
3. 单击**创建集群**，在**基础配置**页面的**高级设置**区域，添加步骤二：创建RAM角色中创建的角色名称。创建详情请参见创建集群。

4. 集群创建成功后，通过SSH登录主节点，详情请参见登录集群。

   执行以下命令，验证授权是否成功。

   ```
   hdfs dfs -ls oss://<yourBucketName>/
   ```

   > **说明** 示例中的<yourBucketName>为您OSS Bucket的名称。

   - 没有该Bucket访问权限时，提示如下信息。

     ```
     [root@emr-header-1 ~]# hdfs dfs -ls oss://emr-logs2/
     ls: java.io.IOException: ErrorCode : 403 , ErrorMsg: HTTP/1.1 403 Forbidden: <?xml version="1.0" enco
     ding="UTF-8"?>
     ```

   - 有该Bucket访问权限时，提示如下信息。

     ```
     [root@emr-header-1 ~]# hdfs dfs -ls oss://emr-logs2/
     Found 10 items
     -rw-rw-rw-   1         30 2020-03-10 16:08 oss://emr-logs2/
     drwxrwxrwx   -          0 2019-12-20 14:49 oss://emr-logs2/
     drwxrwxrwx   -          0 2020-03-12 23:56 oss://emr-logs2/
     -rw-rw-rw-   1        161 2020-05-18 22:01 oss://emr-logs2/
     drwxrwxrwx   -          0 2019-12-20 14:22 oss://emr-logs2/
     drwxrwxrwx   -          0 2020-05-18 22:28 oss://emr-logs2/
     ```

常见问题

- Q：创建集群时提示NoPermission。

  A：您可以参照如下方式排查解决。

  i. 您创建集群使用的用户是否有创建集群和更换ECS应用角色的权限，如果该RAM用户权限为AliyunEMRDevelopAccess可以修改为AliyunEMRFullAccess。

  ii. 创建集群时ECS应用角色名称是否填写正确。

  iii. 授信策略是否修改为*emr.aliyuncs.com*。

- Q：HDFS无法访问OSS路径

  ```
  [root@emr-header-1 ~]# hdfs dfs -ls oss://
  ls: java.io.IOException: ErrorCode : 403 , ErrorMsg: HTTP/1.1 403 Forbidden ERROR_CODE : 1010
  ```

  A：您可以参照如下方式排查解决。

  i. 确认访问的OSS Bucket是否和集群在一个地域（Region），如果不在同一地域（Region），在访问链接中需要添加相应的Endpoint。

  ii. 确认访问的OSS Bucket是否包含在新建的权限策略中，如果没有，需要修改权限策略。

  iii. 确认是否在OSS控制台上设置了该Bucket的相关权限。如果设置了相关权限，您可以在OSS控制台上取消相关权限的设置，通过设置权限策略中的Action内容来设置相关权限。

3.3.6. 为RAM用户授权

为确保RAM用户能正常使用E-MapReduce控制台功能，您需要使用阿里云账号登录访问控制RAM（Resource Access Management）控制台，授予RAM用户相应的权限。

背景信息

访问控制RAM是阿里云提供的资源访问控制服务，详情请参见什么是访问控制。在E-MapReduce中，RAM的典型使用场景如下：

- 用户：如果您购买了多台E-MapReduce集群实例，您的组织里有多个用户（如运维、开发或数据分析）需要使用这些实例，您可以创建一个策略允许部分用户使用这些实例。避免了将同一个AccessKey泄露给多人。
- 用户组：您可以创建多个用户组，并授予不同权限策略，授权过程与授权用户过程相同，实现批量管理用户权限。

权限策略

权限策略分为系统策略和自定义策略：

- 系统策略
  阿里云提供多种具有不同管理目的的系统策略，E-MapReduce使用的系统策略如下。

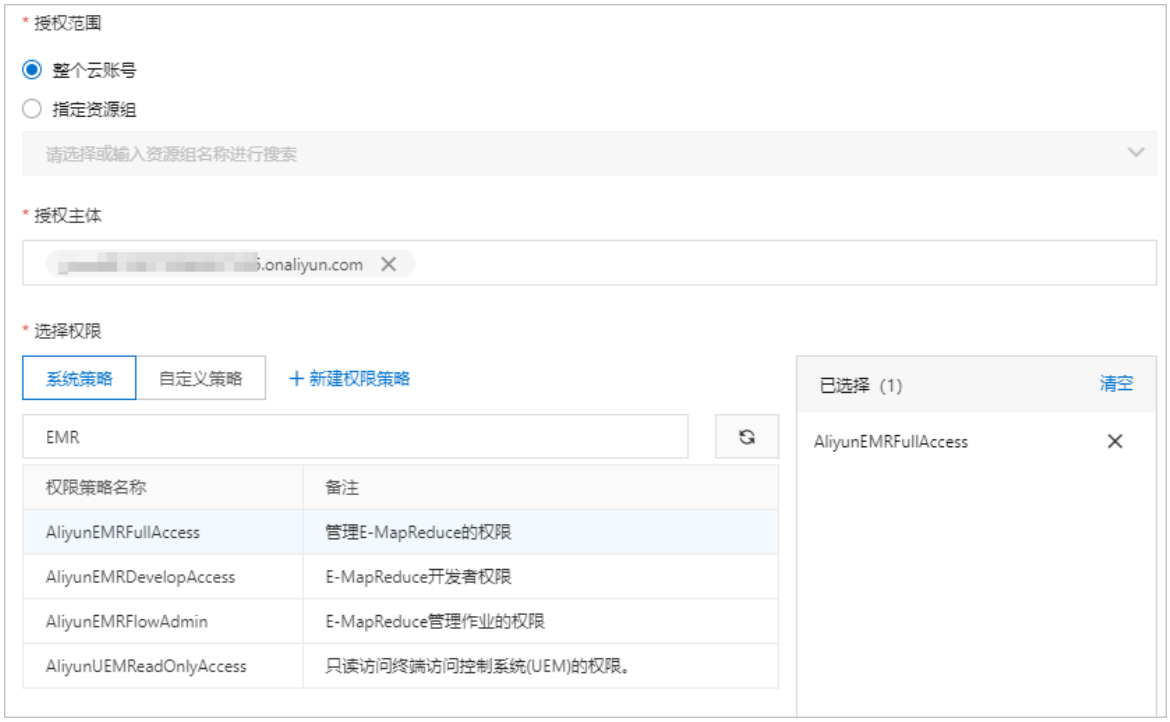
系统策略名称	描述	包含的权限
AliyunEMRFullAccess	E-MapReduce管理员权限	所有权限。
AliyunEMRDevelopAccess	E-MapReduce开发者权限	除集群的创建和释放等权限外的所有权限。
AliyunEMRFlowAdmin	E-MapReduce数据开发的管理员权限	创建项目、开发和管理作业权限（不包含添加项目成员和管理集群权限）。

- 自定义策略
  如果您熟悉阿里云各种云服务API，并且需要精细化的权限控制策略，可参见权限策略语法和结构，创建自定义策略。创建时，您需要精准地设计权限策略脚本。

授权RAM用户

执行以下步骤，在访问控制RAM控制台，为RAM用户授予E-MapReduce相关权限：

1. 使用阿里云账号登录RAM控制台。
2. 在左侧导航栏，选择**身份管理 > 用户**。
3. 单击待授权RAM用户所在行的**添加权限**。
4. 在**添加权限**页面，根据需求选择授权范围、授权主体和权限。



参数	描述
授权范围	○ **整个云账号**：权限在当前阿里云账号内生效。 ○ **指定资源组**：权限在指定的资源组内生效。
授权主体	需要授权的RAM用户。
选择权限	在**系统策略**中，输入EMR搜索EMR相关权限策略，然后单击需要授予RAM用户的权限策略，选择该权限。各权限策略的详细说明，请参见权限策略。

5. 单击**确定**。
6. 单击**完成**。
   完成授权后，权限立即生效，被授权的RAM用户可以登录RAM控制台，执行已被授权的操作。

3.4. 管理用户

本文为您介绍如何通过E-MapReduce（简称EMR）的用户管理功能，管理集群中的EMR用户。

背景信息

EMR用户信息存储在集群自带的OpenLDAP中，主要用于E-MapReduce集群内的身份认证。

EMR用户可以用于访问链接与端口，查看开源组件Web UI时的用户身份认证，也可以在开启组件LDAP认证之后进行身份认证。如果将Ranger的用户源设置为LDAP，则可以对用户管理中的用户进行权限控制。如果是高安全集群，EMR用户可以用于Kinit操作。

用户管理中的EMR用户以列表的形式进行展示和操作。根据登录EMR控制台的RAM用户权限的不同，可以将使用E-MapReduce用户管理模块的RAM用户分为两类：

- 管理员：阿里云账号或者拥有emr:ManageUserPlatform（例如系统策略AliyunEMRFullAccess）和emr:CreateLdapUser权限的RAM用户。管理员可以查看集群中所有用户列表，并对所有用户执行重置密码、删除、修改备注、下载认证凭据（高安全集群）操作，也可以添加用户。

- 普通用户：拥有其他权限策略（例如系统策略AliyunEMRDevelopAccess）的RAM用户。普通用户只能在用户列表中查看与自己同名的EMR用户，并只能进行重置密码、修改备注和下载认证凭据（高安全集群）操作，不能添加和删除用户。

前提条件

- 已创建集群，详情请参见创建集群。
- 已创建RAM用户，详情请参见创建RAM用户。

> **说明** 因为在E-MapReduce用户管理中添加的用户需要与RAM用户同名，因此需要先创建RAM用户。

使用限制

仅Hadoop集群、Data Science集群、Flink集群、EMR Studio集群、Presto集群和Druid集群，支持用户管理功能。

添加用户

> **注意** 如果当前用户使用的是RAM用户，则添加用户时需要该RAM用户拥有ram:ListUsers的权限。您可以使用阿里云账号在访问控制台上给该RAM用户添加AliyunRAMReadOnlyAccess系统策略，或者自定义权限策略并添加ram:ListUsers权限。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏，单击**用户管理**。
3. 在**用户管理**页面，单击**添加用户**。
4. 在**添加用户**对话框的**用户名**下拉列表中，选择已有的RAM用户作为EMR用户的名称，输入**密码**和**确认密码**。
5. 单击**确定**。

删除用户

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏，单击**用户管理**。
3. 在**用户管理**页面，单击目标用户所在行的**删除**。
4. 在**删除**对话框中，单击**确定**。

重置用户密码

您可以修改已添加用户的密码。

> **注意** 重置密码可能导致正在运行的任务失败。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏，单击**用户管理**。
3. 在**用户管理**页面，单击目标用户所在行的**重置密码**。
4. 在**设置密码**对话框中，输入新的**密码**和**确认密码**。
5. 单击**确定**。

下载认证凭据

> **注意** 下载认证凭据功能仅支持开启高安全的集群，通过该功能，您可以下载目标用户的Keytab文件。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏，单击**用户管理**。
3. 在**用户管理**页面，单击目标用户所在行的**下载认证凭据**。

更新用户

用于刷新由于网络延迟导致的未及时生效的用户管理操作，也可以用于将直接添加的OpenLDAP用户同步至用户管理中。

在**用户管理**页面，单击**更新**，更新用户。

管理Linux用户

主要面向自建LDAP且开启高安全的集群的场景。添加Linux用户时，EMR会自动为每台主机创建指定名称的Linux用户，此后扩容的机器上也会有这个用户。

1. 在**用户管理**页面，单击右上角的**Linux用户管理**。
2. 在**添加用户**输入框中，输入用户名称，多个用户通过分号（;）分隔。
3. 单击**添加**。

常见问题

Q：不同集群中的用户是互通的吗？

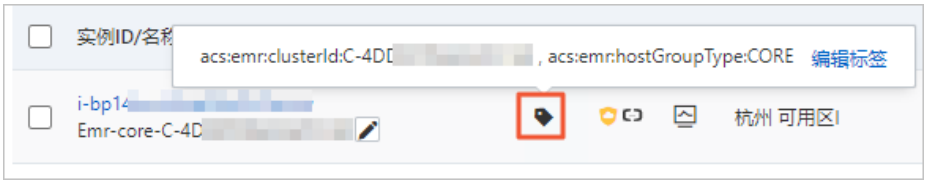
A：不互通。因为用户管理中的用户只在当前集群中生效，所以每个集群的EMR用户不互通。例如，在cluster-1上添加EMR用户A之后，并不会共享给cluster-2。如果需要在cluster-2上使用EMR用户A，则需要在cluster-2上重新添加用户A。

3.5. 设置标签

标签是集群的标识。为集群添加标签，可以方便您识别和管理拥有的集群资源。您可以在创建集群时绑定标签，也可以在集群创建完成后，在集群基础信息页面绑定标签，您最多可以给集群绑定二十个标签。本文为您介绍如何新建、查看和绑定标签等。

背景信息

E-MapReduce集群中包含多个ECS节点。创建E-MapReduce集群时，系统会自动为ECS节点绑定两个系统标签。通过登录云服务器ECS控制台，在实例列表中将鼠标移动到对应ECS节点的标签图标上，可以查看该ECS节点属于哪个集群以及在集群中的角色。



例如，某个ECS节点上的系统标签如下：

- acs:emr:clusterId=C-A510C93EA117****
- acs:emr:hostGroupType=CORE

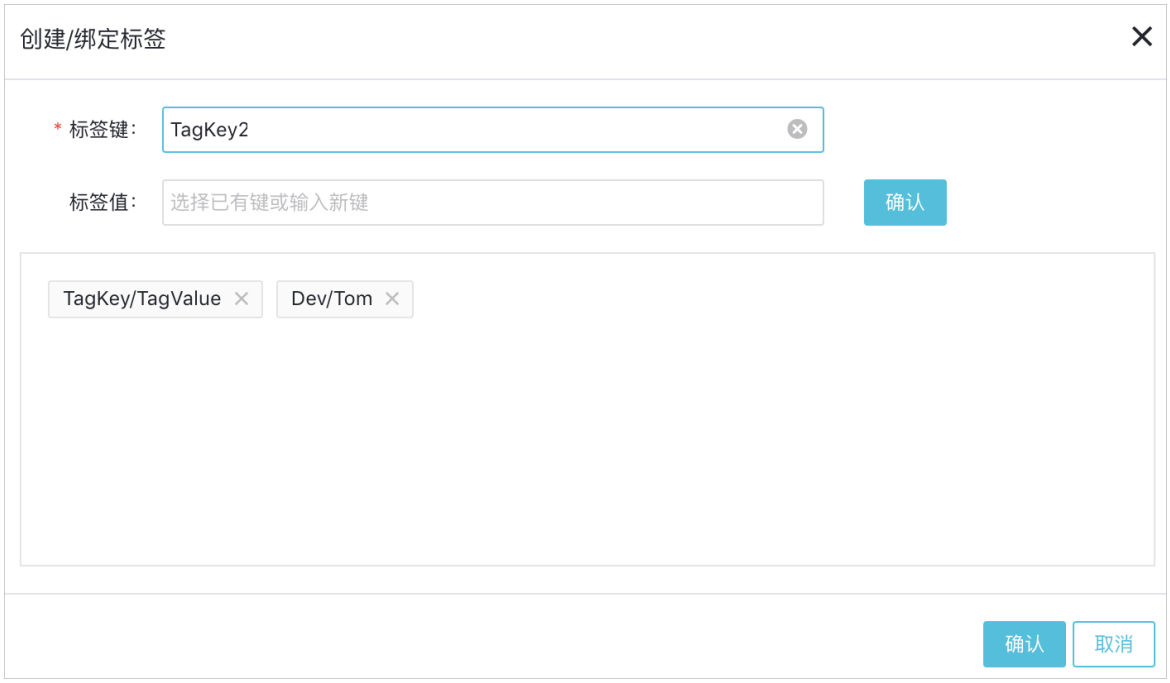
表示该节点是集群ID为*C-A510C93EA117****的EMR集群下的一个CORE节点。

注意事项

- 更新E-MapReduce的标签会同步到对应的ECS节点上。
- 更新ECS节点的标签不会同步到集群E-MapReduce上，因此为了保持ECS节点与E-MapReduce上标签的一致性，建议不要单独在ECS控制台上修改ECS的标签。并且当集群中某个ECS节点的标签数量达到上限时，集群将不能再创建标签。
- 不同地域中的标签信息不互通。
  例如，在华东 1（杭州）地域创建的标签在华东 2（上海）地域不可见。
- 一个资源绑定标签的上限为20个。如果超出上限，您需要解绑部分标签后再继续绑定新标签。

为集群创建或绑定标签

1. 登录阿里云E-MapReduce控制台。
2. 在顶部菜单栏处，根据实际情况选择地域和资源组。
3. 单击**创建集群**，在**基础配置**页面的**高级设置**区域，单击**标签**所在行的**添加**。
   创建集群详情请参见创建集群。
4. 在**创建/绑定标签**对话框中，选择已有标签键或输入新标签键。
   选择已有标签键和标签值意味着绑定标签，输入新标签键意味着创建标签。



> **说明** 标签都是由一对键值对组成，标签键和标签值需满足以下约束信息：
> - 标签键不可以为空，最长为128位；标签值可以为空，最长为128位。
> - 标签键和标签值不能以aliyun或acs:开头，不能包含http://和https://。
> - 任一标签的标签键（Key）必须唯一。例如，集群先绑定了city/shanghai，后续如果绑定city/newyork，则city/shanghai自动被解绑。

5. 单击**确认**。

   确认待绑定标签。

6. 单击**确认**。

   完成绑定标签。

使用标签搜索集群

在集群管理页，按标签键或标签值搜索目标集群。

1. 进入集群管理页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
2. 在**集群管理**页面，单击**标签**，选择一个标签键。

   如果您未选择具体的标签值，默认展示该标签键绑定的所有集群。
3. 单击**标签**所在行的**搜索**。

   下方列表会为您展示满足搜索条件的集群。

管理已有集群的标签

对现有集群的标签进行绑定、解绑、查看操作。

1. 进入详情页面。

i. 登录阿里云E-MapReduce控制台。
ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
iii. 单击上方的**集群管理**页签。
iv. 在集群列表中，单击待操作集群所在行的**详情**。

2. 在**集群基础信息**页面，单击**集群信息**区域的**编辑标签**。
3. 在**编辑标签**页面，管理目标集群的标签。
   - 创建新标签。
     a. 单击下方的**新建标签**。
     b. 输入**标签键**和**标签值**。
     c. 单击**确认**。
     d. 单击**确认**。
   - 绑定已有标签。
     a. 单击下方的**已有标签**。
     b. 选择已有**标签键**和**标签值**。
     c. 单击**确认**。
   - 解绑标签。
     单击标签栏中每条标签后的 × 图标，解绑该标签。

> **说明** 解绑标签时，如果解绑之后该标签不再绑定任何资源，则该标签会自动被删除。

3.6. 使用资源组

资源组会对您拥有的云资源从用途、权限和归属等维度上进行分组，实现企业内部多用户、多项目的资源分级管理。一个云资源只能属于一个资源组，云资源之间的关联关系不会因加入资源组而发生变化。E-MapReduce目前支持资源组的云资源为集群和项目。本文介绍如何为云资源指定资源组及相关的示例。

背景信息

使用资源组时，您需要了解：

- 一个资源组可以包含不同地域的云资源。例如，资源组A中可以包含华东1（杭州）地域的集群和华东2（上海）地域的集群。
- 一个资源组可以包含不同资源类型。例如，资源组A中可以包含集群、ECS实例和项目等多种云资源。
- 同一个账号内不同资源组中，相同地域的集群和项目可以进行关联。如果当前账号同时具有资源组A和资源组B的相关权限，则资源组A中华北2（北京）地域项目的作业可以运行在资源组B中华北2（北京）地域的集群上。
- 资源组会继承RAM用户的全局权限。即如果您授权RAM用户管理所有的阿里云资源，那么阿里云账号下所有的资源组都会在该RAM用户中显示出来。

使用限制

- 资源组的创建、管理和RAM授权都是在阿里云资源管理（Resource Management）控制台上进行，详情请参见什么是资源管理。
- E-MapReduce支持资源组的云资源为集群和项目。在创建、扩容集群或转移集群资源组时，集群每个节点会同步加入集群所属资源组，节点支持资源组的云资源包括ECS实例、云盘、镜像、弹性网卡、安全组和密钥对。

- 不允许跨账号在资源组之间转移云资源。

> **注意** 转移节点相关资源的资源组不会同步到集群上，为了保持集群相关资源的统一管理和RAM授权，不建议在资源管理控制台上单独对节点相关资源（例如，ECS实例、云盘、镜像、弹性网卡、安全组和密钥对）转移所属资源组。

指定资源组

云资源在创建之后必须属于一个资源组，如果在创建时未指定，则会加入默认的资源组。下面介绍如何在创建集群和项目时指定资源组。

> **说明** 每个账号在创建云资源时只能将云资源加入有相关权限的资源组中。

- 创建集群
  i. 登录阿里云E-MapReduce控制台。
  ii. 在顶部菜单栏处，选择地域（Region）。
  iii. 单击**创建集群**。
  iv. 完成软硬件配置后，在**基础配置**页面，展开**高级设置**，在**资源组**区域中选择已有的资源组。
  如果需要创建新的资源组，您可以单击**去创建**来新建资源组，详情请参见创建资源组。

  > **说明** 创建集群的详细信息请参见创建集群。

- 创建项目
  i. 登录阿里云E-MapReduce控制台。
  ii. 在顶部菜单栏处，选择地域（Region）。
  iii. 单击上方的**数据开发**页签。
  iv. 在**数据开发**页面，单击右上角的**新建项目**。
  v. 在**新建项目**对话框中，输入**项目名称**和**项目描述**，从**资源组选择**列表，选择已有的资源组。
  如果需要创建新的资源组，您可以单击**去创建**来新建资源组，详情请参见创建资源组。
  vi. 单击**创建**。
  创建项目的详细信息请参见项目管理。

应用场景

资源组有以下两个应用场景：

- 将不同用途的云资源加入不同的资源组中分开管理。详情请参见场景一：按云资源的用途分组。
- 为各个资源组设置完全独立的管理员，实现资源组范围内的用户与权限管理。详情请参见场景二：资源组范围内的用户与权限管理。

> **说明** 由于授权原因RAM用户仅能看到被授权的资源，因此对于只有部分资源组权限而没有全局权限的RAM用户，在顶部菜单栏选择**账号全部资源**时会报无权限的提示。

场景一：按云资源的用途分组

您可以将生产环境和测试环境的集群，分别放入生产环境和测试环境两个资源组中。产品测试时，建议只对测试环境资源组中的集群进行操作，避免对生产环境的集群发生误操作。当产品需要上线时，再选择生产环境资源组中的集群进行操作。

1. 创建资源组**测试环境**和**生产环境**。

详情请参见创建资源组。

2. 为资源组**测试环境**和**生产环境**设置同一个管理员。

   详情请参见添加RAM身份并授权。

3. 创建集群TestEnv1和TestEnv2。

   创建集群时，请指定加入**测试环境**资源组。

4. 创建集群ProdEnv1和ProdEnv2。

   创建集群时，请指定加入**生产环境**资源组。

5. 使用**测试环境**和**生产环境**资源组的管理员账号登录阿里云E-MapReduce控制台。

6. 在顶部菜单栏处选择相应资源组。

   您可以看到相应资源组的集群都显示在集群列表页面。例如，选择**测试环境**，您只可以看到测试环境的集群显示在集群列表页面。

场景二：资源组范围内的用户与权限管理

您可以将公司不同部门使用的集群和项目分别放入多个资源组中，并设置相应的管理员分部门管理集群和项目，实现资源组的隔离。本示例以**开发部**和**测试部**进行介绍。

1. 创建资源组**开发部**和**测试部**。

   详情请参见创建资源组。

2. 分别为资源组**开发部**和**测试部**设置管理员。

   详情请参见添加RAM身份并授权。

3. 创建集群ITCluster和项目ITFlowProject，在创建时指定**开发部**资源组。

   集群创建的详细信息请参见创建集群；项目创建的详细信息请参见项目管理。

4. 创建集群FinanceCluster1和FinanceCluster2。在创建时指定**测试部**资源组。

5. 使用测试部管理员账号登录阿里云E-MapReduce控制台。

6. 在顶部菜单栏处选择**测试部**。

   您可以看到属于测试部的集群都显示在集群列表页面。

3.7. 管理安全组

安全组用于设置集群内ECS实例的网络访问控制，是重要的安全隔离手段。本文为您介绍如何添加安全组及安全组规则。

背景信息

创建集群时，您必须选择或新建一个安全组，对某个安全组下的所有ECS实例的出方向和入方向进行网络控制。

您可以将ECS实例按照功能划分，放于不同的安全组中。例如，通过E-MapReduce创建的安全组为E-MapReduce安全组，而您已有的安全组为用户安全组，每个安全组按照不同的需要设置不同的访问控制策略。

注意事项

- 设置安全组规则时，一定要限制访问IP范围，且不要使用0.0.0.0/0，避免被攻击。
- 在添加安全组规则的时候，开放应用出入规则时应遵循最小授权原则，授权对象只针对当前服务器的公网访问IP地址开放。
  您可以通过访问IP地址，获取当前服务器的公网访问IP地址。

创建安全组

您在创建E-MapReduce集群时，可以新建安全组或者使用已有的安全组。

> **注意** 禁止使用ECS上创建的企业安全组。

添加安全组

> **说明**
> - 经典网络类型下，实例必须加入同一地域下经典网络类型的安全组。
> - 专有网络类型下，实例必须加入同一专有网络下的安全组。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在**主机信息**区域，单击**主实例组 (MASTER)** 或**核心实例组 (CORE)**，再单击右侧某个ECS实例的**ECS ID**。
3. 在该ECS实例页面，单击**安全组**页签。
4. 在**安全组列表**页面，单击**加入安全组**。
5. 在**ECS实例加入安全组**对话框中，从**安全组**列表中选择需要加入的安全组。

   如果您需要将该ECS实例一次加入多个安全组，选择一个安全组后，单击后面的**加入到批量选择栏**，即会把该安全组加入到批量选择栏中，依次按照相同方法再选择其他安全组，把其他安全组也加入到批量选择栏中即可。
6. 单击**确定**。

   请重复步骤2~步骤6操作，直至把E-MapReduce集群中的其他ECS实例也加入到相应安全组中。

添加安全组规则

1. 获取机器的公网访问IP地址。

   为了安全地访问集群组件，在设置安全组策略时，推荐您只针对当前的公网访问IP地址开放。您可以通过访问IP地址，获取当前服务器的公网访问IP地址。
2. 进入EMR的**详情**页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
   v. 在**集群基础信息**页面的**网络信息**区域，单击**安全组ID**链接。
3. 在**安全组规则**页面，填写安全组策略。

   修改**端口范围**和**授权对象**，其余参数保持默认。详细信息，请参见添加安全组规则。

参数	说明
端口范围	填写允许访问该ECS实例的端口。
授权对象	填写步骤1中获取的公网访问IP地址。 **注意** 为防止被外部的用户攻击导致安全问题，**授权对象**禁止填写为**0.0.0.0/0**。

4. 单击**保存**。

3.8. 管理元数据

3.8.1. 数据湖元数据

本文介绍如何配置数据湖构建（Data Lake Formation，DLF），作为E-MapReduce（简称EMR）上Hadoop集群的元数据。

背景信息

阿里云数据湖构建是一款全托管的快速帮助用户构建云上数据湖的服务，产品为云原生数据湖提供了统一的元数据管理、统一的权限与安全管理、便捷的数据入湖能力以及一键式数据探索能力，详细信息请参见数据湖构建产品简介。

您可以快速完成云原生数据湖方案的构建与管理，并可无缝对接多种计算引擎，打破数据孤岛，洞察业务价值。

前提条件

已在数据湖构建（Data Lake Formation）控制台开通数据湖构建。

使用限制

- 数据湖元数据适配EMR的Hive 2.x、Hive 3.x、Presto和SparkSQL。
- 仅EMR-3.30.0及之后版本和EMR-4.5.0及之后版本，支持选择数据湖元数据作为Hive数据库。
- 数据湖元数据产品支持华北2（北京）、华东1（上海）、华东2（杭州）和华南1（深圳）地域。

切换元数据存储类型

您可以通过修改Hive参数的方式，切换Hive MetaStore的存储方式。

说明 如果需要迁移数据库的元数据信息，请提交工单处理。

1. 进入Hive服务页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
   v. 在左侧导航栏，选择**集群服务** > **Hive**。
2. 修改**hive.imetastoreclient.factory.class**的值。
   i. 在Hive服务页面，单击**配置**页面。

ii. 在配置搜索中，输入配置项hive.imetastoreclient.factory.class，单击 图标。

- 切换为MySQL（包括集群内置MySQL、统一meta数据库和独立RDS MySQL）：
  设置hive.imetastoreclient.factory.class的值
  为org.apache.hadoop.hive.ql.metadata.SessionHiveMetaStoreClientFactory。
- 切换为数据湖元数据：
  设置hive.imetastoreclient.factory.class的值
  为com.aliyun.datalake.metastore.hive2.DlfMetaStoreClientFactory。

3. 保存配置。
   i. 在Hive服务页面，单击右上角的**保存**。
   ii. 在**确认修改**对话框中，输入执行原因，单击**确定**。
4. 重启Hive MetaStore服务。
   i. 在Hive服务页面，选择右上角的**操作 > 重启Hive MetaStore**。
   ii. 在**执行集群操作**对话框，设置相关参数，然后单击**确定**。
   iii. 在**确认**对话框，单击**确定**。
   您可以单击右上角的**查看操作历史**，查看任务进度，等待任务完成。

3.8.2. 配置独立RDS

本文介绍如何配置独立的阿里云RDS，作为E-MapReduce（简称EMR）上Hadoop或EMR Studio集群的元数据。

背景信息

EMR Studio正处于公测阶段，如果您需要体验，请提交工单。更多EMR Studio信息，请参见EMR Studio概述。

前提条件

已购买RDS MySQL实例，详情请参见创建RDS MySQL实例。

> **说明** 本文以MySQL 5.7版本为例介绍。

使用限制

EMR上创建的集群类型与MySQL关系如下：

- 如果创建的是Hadoop集群，则数据库**类型**选择**MySQL**，版本选择5.7；**系列**选择**高可用版**。
- 如果创建的是EMR Studio集群，则数据库**类型**选择**MySQL**，版本选择8.0。

操作流程

1. 步骤一：元数据库准备
   准备元数据库信息。
2. 步骤二：创建集群
   在EMR控制台上创建集群，关联元数据库信息。
3. （可选）步骤三：Metastore初始化
   根据Hive版本初始化Metastore。

   > **注意** 如果您创建的是Hadoop集群，则需要执行该步骤。

步骤一：元数据库准备

1. 创建hivemeta的数据库。

   详情请参见创建数据库和账号中的创建数据库。

2. 创建用户并授权读写权限。

> 注意
>
> - 如果您创建的是Hadoop集群，则可以创建普通账号，详情请参见创建数据库和账号。
> - 如果您创建的是EMR Studio集群，则需要创建高级权限账号，详情请参见创建数据库和账号。

请记录创建账号的用户名和密码，步骤二：创建集群会用到。

3. 获取数据库内网地址。
   i. 设置白名单，详情请参见通过客户端、命令行连接RDS MySQL。
   ii. 在实例详细页面，单击左侧导航栏中的**数据库连接**。

iii. 在**数据库连接**页面，单击内网地址进行复制。

请记录内网地址，步骤二：创建集群时会用到。

步骤二：创建集群

在创建集群的**基础配置**页面，配置以下参数，其他参数的配置请参见创建集群。

参数	描述
集群名称	集群的名字，长度限制为1~64个字符，仅可使用中文、字母、数字、中划线（-）和下划线（_）。
元数据选择	选择**独立RDS MySQL**。
数据库链接	**数据库连接**填写格式为*jdbc:mysql://rm-xxxxxx.mysql.rds.aliyuncs.com/<数据库名称>?createDatabaseIfNotExist=true&characterEncoding=UTF-8*。 • *rm-xxxxxx.mysql.rds.aliyuncs.com*为步骤一：元数据库准备中获取的数据库内网地址。 • *<数据库名称>*为步骤一：元数据库准备中设置的数据库名称。
数据库用户名	填写步骤一：元数据库准备中账号的用户名。
数据库密码	填写步骤一：元数据库准备中账号的密码。
数据开发存储	设置Airflow的logs、dags以及Zeppelin的notebook在OSS上的存储位置。 **说明** 该参数仅在创建EMR Studio集群时可见。

步骤三：Metastore初始化

注意 如果您创建的是Hadoop集群，则需要按照以下步骤根据Hive版本初始化Metastore。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。

iii. 单击上方的**集群管理**页签。

iv. 在**集群管理**页面，单击相应集群所在行的**详情**。

2. 在**集群基础信息**页面的**软件信息**区域，检查Hive的版本，并进行初始化。

- 如果Hive是2.3.x版本，请按照以下步骤进行初始化。

a. 使用ssh方式登录集群Master节点，详情请参见登录集群。

b. 执行以下命令，进入mysql目录。

```
cd /usr/lib/hive-current/scripts/metastore/upgrade/mysql/
```

c. 执行以下命令，登录MySQL数据库。

```
mysql -h {RDS数据库内网或外网地址} -u{RDS用户名} -p{RDS密码}
```

d. 执行以下命令，进行初始化。

```
use {RDS数据库名称};
source /usr/lib/hive-current/scripts/metastore/upgrade/mysql/hive-schema-2.3.0.mysql.sql;
```

上述命令中参数描述如下：

- `{RDS用户名}` 为步骤一：元数据库准备中账号的用户名。
- `{RDS密码}` 为步骤一：元数据库准备中账号的密码。本示例为hiveuser。
- `{RDS数据库名称}` 为步骤一：元数据库准备中设置的数据库名称。本示例为hivemeta。

- 如果Hive是其他版本时，执行以下命令进行初始化。

a. 使用ssh方式登录集群Master节点，详情请参见登录集群。

b. 执行以下命令，切换为hadoop用户。

```
su hadoop
```

c. 执行以下命令，登录MySQL数据库。

```
schematool -initSchema -dbType mysql
```

待初始化成功后，则可以使用自建的RDS作为Hive的元数据库。

> **说明** 在初始化之前，Hive的Hive MetaStore、HiveServer2和Spark的ThriftServer可能会出现异常，待初始化之后会恢复正常。

常见问题

- Metastore初始化时提示Failed to get schema version异常信息，该如何处理？
- 如果Hive元数据信息中包含中文信息，例如列注释和分区名等，该如何处理？

3.9. 管理引导操作

添加引导操作，可以安装您需要的第三方软件或者修改集群运行环境。本文为您介绍如何添加、编辑、克隆和删除引导操作。

背景信息

通过引导操作，您可以完成很多目前E-MapReduce集群尚未支持的操作，例如：

- 使用Yum安装已经提供的软件。
- 直接下载公网上的一些公开的软件。
- 读取OSS中您的自有数据。
- 安装并运行一个服务，例如Flink或者Impala。

使用限制

- 您最多可以添加10个引导操作。引导操作会按照您指定的顺序执行。
- 您指定的脚本默认使用root账户执行，您也可以在脚本中使用**su hadoop**命令，切换为hadoop用户执行。

添加引导操作

添加引导操作支持以下两种方式：

- 方式一：创建集群时添加引导操作
  i. 进入集群管理页面。
     a. 登录阿里云E-MapReduce控制台。
     b. 在顶部菜单栏处，根据实际情况选择地域和资源组。
     c. 单击上方的**集群管理**页签。
  ii. 单击右上角的**创建集群**。
  iii. 在**基础配置**的**高级设置**区域，单击**引导操作**所在行的**添加**。
  iv. 在**添加引导操作**对话框中，填写配置项。

参数	描述
名称	引导操作的名称。
脚本位置	选择脚本所在OSS的位置。 脚本路径格式必须为*oss://**/*.sh*格式。
参数	引导操作脚本的参数，指定脚本中所引用的变量的值。
执行范围	■ **集群**：此引导操作适用于整个集群。 ■ **机器组**：此引导操作适用于您选择的机器组。 您可以选择主实例、核心实例组或者已有的机器组。
执行时间	■ **组件启动前**：组件启动前执行该脚本。 ■ **组件启动后**：组件启动后执行该脚本。
执行失败策略	■ **继续执行**：如果该脚本执行失败，继续执行下一个脚本。 ■ **停止执行**：如果该脚本执行失败，停止当前脚本。

  v. 单击**确定**。
  引导操作示例请参见示例。

> **说明** 引导操作可能会执行失败，但引导操作失败并不会影响集群的创建。

集群创建成功后，您可以在**集群基础信息**页面，查看**引导操作/软件配置**是否有异常发生。如果有异常，您可以登录到各个节点上查看运行日志，运行日志在*/var/log/bootstrap-actions*目录下。

- 方式二：创建集群后添加引导操作
  i. 进入引导操作页面。
     a. 登录阿里云E-MapReduce控制台。
     b. 在顶部菜单栏处，根据实际情况选择地域和资源组。
     c. 单击上方的**集群管理**页签。
     d. 在**集群管理**页面，单击相应集群所在行的**详情**。
     e. 在左侧导航栏，单击**引导操作**。
  ii. 在**引导操作**页面，单击**添加引导操作**。
  iii. 在**添加引导操作**对话框中，填写配置项。

参数	描述
名称	引导操作的名称。
脚本位置	选择脚本所在OSS的位置。 脚本路径格式必须为*oss://**/*.sh*格式。
参数	引导操作脚本的参数，指定脚本中所引用的变量的值。
执行范围	■ **集群**：此引导操作适用于整个集群。 ■ **机器组**：此引导操作适用于您选择的机器组。 您可以选择主实例、核心实例组或者已有的机器组。
执行时间	■ **组件启动前**：组件启动前执行该脚本。 ■ **组件启动后**：组件启动后执行该脚本。
执行失败策略	■ **继续执行**：如果该脚本执行失败，继续执行下一个脚本。 ■ **停止执行**：如果该脚本执行失败，停止当前脚本。

  iv. 单击**确定**。
  引导操作示例请参见示例。

编辑引导操作

1. 进入引导操作页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
   v. 在左侧导航栏，单击**引导操作**。
2. 在**引导操作**页面，单击待编辑引导操作所在行的**编辑**。
3. 在**编辑引导操作**对话框中，您可以根据需求修改所有配置项。

4. 单击**确定**。

克隆引导操作

1. 进入引导操作页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
   v. 在左侧导航栏，单击**引导操作**。
2. 在**引导操作**页面，单击待克隆引导操作所在行的**克隆**。
3. 在**克隆引导操作**对话框中，您可以根据需求修改所有配置项。
4. 单击**确定**。

删除引导操作

1. 进入引导操作页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
   v. 在左侧导航栏，单击**引导操作**。
2. 在**引导操作**页面，单击待删除引导操作所在行的**删除**。
3. 在**删除引导操作**对话框中，单击**确认**。

示例

引导操作需要指定引导操作的名称和执行脚本在OSS上的位置，根据需要指定脚本的参数。执行引导操作时各个节点会下载您指定的OSS脚本，直接执行或者附加上可选参数执行。引导操作示例如下：

- 示例1
  您可以在脚本中指定需要从OSS下载的文件。例如，添加以下脚本，将*oss://<yourbucket>/<myfile>.tar.gz*文件下载到本地，并解压到*/<yourdir>*目录下：

  > **注意** OSS地址有内网地址、外网地址和VPC网络地址之分。如果使用经典网络，则需要指定内网地址，例如杭州是*oss-cn-hangzhou-internal.aliyuncs.com*。如果使用VPC网络，则需要指定VPC内网可访问的域名，例如杭州是*vpc100-oss-cn-hangzhou.aliyuncs.com*。

  ```
  #!/bin/bash
  osscmd --id=<yourid> --key=<yourkey> --host=oss-cn-hangzhou-internal.aliyuncs.com get oss:/
  /<yourbucket>/<myfile>.tar.gz ./<myfile>.tar.gz
  mkdir -p /<yourdir>
  tar -zxvf <myfile>.tar.gz -C /<yourdir>
  ```

- 示例2
  您可以通过Yum安装额外的系统软件包，例如安装*ld-linux.so.2*：

  ```
  #!/bin/bash
  yum install -y ld-linux.so.2
  ```

3.10. 集群脚本

集群创建完成后（特别是包年包月集群），您可以通过集群脚本功能批量选择节点来运行指定脚本，以实现个性化需求。例如，安装第三方软件和修改集群运行环境。

背景信息

集群脚本适用于长期存在的集群，对按需创建的临时集群，应使用引导操作来完成集群初始化工作。集群脚本类似引导操作，在集群创建完成后，您可以通过集群脚本功能来安装集群尚未支持的软件和服务，例如：

- 使用YUM安装已经提供的软件。
- 直接下载公网上公开的软件。
- 读取您OSS中的自有数据。
- 安装并运行服务（例如，Flink或者Impala），但需要编写的脚本会复杂些。

前提条件

- 已创建集群，详情请参见创建集群。
- 请确保集群状态是空闲或运行中，其他状态时集群不支持运行集群脚本。
- 已开发或已获取集群脚本（示例），并上传到OSS。

注意事项

- 一个集群同一时间只能运行一个集群脚本，如果有正在运行的集群脚本，则无法再提交执行新的集群脚本。每个集群最多保留10个集群脚本记录，如果超过10个，则需要您将之前的记录删除才能创建新的集群脚本。
- 集群脚本可能在部分节点上运行成功，部分节点上运行失败。例如，节点重启导致脚本运行失败。在解决异常问题后，您可以单独指定失败的节点再次运行。当集群扩容后，您也可以指定扩容的节点单独运行集群脚本。
- 集群脚本会在您指定的节点上下载OSS上的脚本并运行，如果运行状态是失败，则您可以登录到各个节点上查看运行日志，运行日志存储在每个节点的*/var/log/cluster-scripts/clusterScriptId*目录下。如果集群配置了OSS日志目录，运行日志也会上传到*osslogpath/clusterId/ip/cluster-scripts/clusterScriptId*目录。

使用集群脚本

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧的导航栏，单击**集群脚本**。
3. 在**集群脚本**页面，单击右上角的**创建并执行**。
4. 在**创建脚本**对话框中，输入**名称**，选择**脚本**，选中**执行节点**。

   > **说明** 在使用集群脚本功能时，强烈建议您先在单个节点对集群脚本进行测试，在测试全部正常后，再在整个集群上操作。

5. 完成上述参数配置后，单击**确定**。

   集群脚本创建完成后，会显示在集群脚本列表中，并且脚本处于运行中状态。

   - 单击**刷新**，可以更新集群脚本的状态。

- 单击**详情**，可以查看脚本在各个节点上的运行情况。
- 单击**删除**，可以删除创建的集群脚本。

示例

与引导操作的脚本相似，您可以在集群脚本中指定需要从OSS下载的文件，以下示例下载*oss://yourbucket/myfile.tar.gz*文件到本地，并解压到*/yourdir*目录下。

```
#!/bin/bash
osscmd --id=<yourAccessKeyId> --key=<yourAccessKeySecret> --host=oss-cn-hangzhou-internal.ali
yuncs.com get oss://<yourBucketName>/<yourFile>.tar.gz ./<yourFile>.tar.gz
mkdir -p /<yourDir>
tar -zxvf <yourFile>.tar.gz -C /<yourDir>
```

> **说明** OSS地址有内网地址、外网地址和VPC网络地址之分。如果是经典网络，则需要指定内网地址（例如，杭州是*oss-cn-hangzhou-internal.aliyuncs.com*）。如果是VPC网络，则需要指定VPC内网可以访问的域名（例如，杭州是*vpc100-oss-cn-hangzhou.aliyuncs.com*）。

集群脚本也可以通过YUM安装额外的系统软件包。例如安装*ld-linux.so.2*。

```
#!/bin/bash
yum install -y ld-linux.so.2
```

集群默认使用root账户来执行您指定的脚本。您也可以在脚本中使用 `su hadoop` 切换到hadoop账户。

3.11. 管理集群资源

3.11.1. 集群资源概述

E-MapReduce（简称EMR） 集群资源管理主要应用于大集群多租户场景中。目前仅支持对E-MapReduce Hadoop类型的集群进行管理。

背景信息

管理EMR集群资源可以帮助您实现以下目标：

- 集群资源中不同部门或用户使用不同的资源队列，实现队列资源的隔离。
- 各队列具有一定的弹性，提高集群的使用效率。

资源管理

EMR集群资源管理目前支持Capacity Scheduler和Fair Scheduler两种调度器。

- Capacity Scheduler：详细操作请参见Capacity Scheduler使用说明。
- Fair Scheduler：详细操作请参见Fair Scheduler使用说明。

> **说明** 官网更多信息请参见Capacity Scheduler和Fair Scheduler。

3.11.2. Capacity Scheduler使用说明

Capacity Scheduler

Capacity Scheduler称为容量调度器，是Apache YARN内置的调度器，E-MapReduce YARN使用Capacity Scheduler作为默认调度器。Capacity Scheduler是一种多租户、分层级的资源调度器，调度器中的子队列是通过设置Capacity来划分各个子队列的使用情况。

前提条件

已创建Hadoop集群，详情请参见创建集群。

注意事项

开启集群的Capacity Scheduler资源队列后，YARN组件**配置**页面中的**capacity-scheduler**配置区域将处于冻结状态，相关已有配置将会同步到**集群资源管理**页面中。如果需要继续在YARN服务的**配置**页面通过XML的方式设置集群资源，则需先在**集群资源管理**中关闭YARN资源队列。

配置Capacity Scheduler

1. 进入集群资源管理页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
   v. 在左侧导航栏中，单击**集群资源管理**。
2. 开启Capacity Scheduler。
   i. 在**集群资源管理**页面，打开**开启YARN资源队列**开关。
   ii. 单击**Capacity Scheduler**。

   > **说明** 首次使用资源管理功能时，资源队列类型默认为Capacity Scheduler。

   iii. 单击**保存**。
3. 在**集群资源管理**页面，单击上方的**资源队列配置**。
4. 在**队列配置**区域，配置队列信息。
   - 单击**编辑**，可以更改资源队列。
   - 单击**更多 > 创建子队列**，可以创建子队列。
     default队列无法创建子队列。

   root为一级队列，是所有队列的父队列，管理YARN的所有资源，默认root下只有default队列。

   > **注意**
   > - 同一个父队列下的同一级别的子队列Capacity之和必须为100。例如，root下有两个子队列default和department，两个子队列的Capacity求和必须为100，department下又设置了market和dev，这里department下两个子队列之和必须也是100。
   > - 在程序运行时，如果不指定提交队列，默认会提交到default队列。
   > - 新建root下一级队列无需重启ResourceManager，直接单击**部署生效**即可。
   > - 新建和设置root下的二级队列需要重启ResourceManager。
   > - 修改队列名称需要重启ResourceManager。

切换调度器类型

开启YARN资源队列后，可以执行以下步骤切换调度器类型。

> **注意** 切换资源调度器后，需要重启服务，才能生效。

1. 在**集群资源管理**页面，单击上方的**切换调度器**。
2. 单击需要切换的资源队列类型。
3. 单击**保存**。
4. 重启服务。
   i. 在**集群资源管理**页面，选择**操作 > 重启ResourceManager**。
   ii. 在**执行集群操作**对话框中，设置相关参数，单击**确定**。
   iii. 在**确认**对话框中，单击**确定**。
   待提示操作成功时，表示切换调度器类型成功。

关闭资源队列

> **说明** 关闭资源队列后，不能在资源管理页面进行任何操作，如果您需要再次配置资源，可以开启资源队列或者在YARN组件**配置**页面中的capacity-scheduler配置区域进行操作 。

1. 在集群资源管理页面，关闭**开启YARN资源队列**开关。
2. 在**关闭资源队列**对话框中，单击**确定**。

提交作业

- 提交作业时，如果不指定提交队列，会默认提交到default队列。
- 指定队列时应选择子节点，不能将任务提交到父队列。
- 提交到指定队列时需通过**mapreduce.job.queuename**参数来指定队列，示例如下。

```
`hadoop jar /usr/lib/hadoop-current/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.8.5.
jar pi -Dmapreduce.job.queuename=test  2 2`
```

3.11.3. Fair Scheduler使用说明

Fair Scheduler
Fair Scheduler称为公平调度器，是Apache YARN内置的调度器。公平调度器主要目标是实现YARN上运行的应用能公平的分配到资源，其中各个队列使用的资源根据设置的权重（weight）来实现资源的公平分配。

前提条件

已创建Hadoop集群，详情请参见创建集群。

注意事项

开启集群的Fair Scheduler资源队列后，YARN组件**配置**中的**fair-scheduler**配置区域将处于冻结状态，相关已有配置将会同步到**集群资源管理**页面中。如果需要继续在YARN服务的**配置**页面通过XML的方式设置集群资源，则需先在**集群资源管理**中关闭YARN资源队列。

配置Fair Scheduler

1. 进入集群资源管理页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。

iv. 在**集群管理**页面，单击相应集群所在行的**详情**。

v. 在左侧导航栏中，单击**集群资源管理**。

2. 开启Fair Scheduler。

i. 在**集群资源管理**页面，打开**开启YARN资源队列**开关。

ii. 单击**Fair Scheduler**。

> 说明 首次使用资源管理功能时，资源队列类型默认为Capacity Scheduler。

iii. 单击**保存**。

3. 在**集群资源管理**页面，单击上方的**资源队列配置**。

4. 在**队列配置**区域，配置队列信息。

- 单击**编辑**，可以更改资源队列。
- 单击**更多 > 创建子队列**，可以创建子队列。
  default队列无法创建子队列。

root为一级队列，是所有队列的父队列，管理YARN的所有资源，默认root下只有default队列。

> 注意 因为队列配置是嵌套的，所以父队列的参数优先级会高于子队列。当子队列设置的资源用量大于父队列时，在调度器分配资源时，会以父队列的参数设置为限制。
>
> - 设置新队列名称，**资源队列名称**是必填项，队列名称中不能包含英文句号（.）。
> - 设置调度器权重，**权重**是必填项。调度器给队列分配资源时，达到设置的权重时，会认为达到公平状态。该权重同时在各个二级、三级以及更深级别的队列生效，如在同一父队列下面的两个子队列权重分别为2和1，则在2个子队列中都有任务运行时，达到2:1的资源分配比例，会认为达到了公平状态。
> - **最大资源量**表示该队列能被分配到的最大资源量，应该大于**最小资源量**，小于集群YARN所能控制的资源规模。如果设置的大于集群资源规模，生效值是YARN所能控制的资源规模。例如，YARN管理的VCore为16个，但最大资源量设置为20，实际生效为16。
> - 在程序运行时，如果不指定提交队列，默认会提交到default队列。
> - 修改子队列名称，如果修改子节点队列名称后，没有重启ResourceManager，任务仍然能提交到原队列名称，但EMR控制台队列配置不会再显示。重启ResourceManager后，原队列将不再存在。
> - 删除队列时，如果删除的不是root的子队列（即非二级队列），删除后单击**部署生效**即可；如果删除的是root的子队列，需要选择**操作 > 重启ResourceManager**，修改才能生效。

切换调度器类型

开启YARN资源队列后，可以执行以下步骤切换调度器类型。

> 注意 切换资源调度器后，需要重启服务，才能生效。

1. 在**集群资源管理**页面，单击上方的**切换调度器**。
2. 单击需要切换的资源队列类型。
3. 单击**保存**。
4. 重启服务。

i. 在**集群资源管理**页面，选择**操作 > 重启ResourceManager**。

ii. 在**执行集群操作**对话框中，设置相关参数，单击**确定**。

iii. 在**确认**对话框中，单击**确定**。

待提示操作成功时，表示切换调度器类型成功。

关闭资源队列

> **说明** 关闭资源队列后，不能在资源管理页面进行任何操作，如果您需要再次配置资源，可以开启资源队列或者在YARN组件**配置**页面中的fair-scheduler配置区域进行操作 。

1. 在集群资源管理页面，关闭**开启YARN资源队列**开关。
2. 在**关闭资源队列**对话框中，单击**确定**。

提交作业

提交到指定队列时需通过mapreduce.job.queuename参数来指定队列，示例如下。

```
`hadoop jar /usr/lib/hadoop-current/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.8.5.jar pi -Dmapreduce.job.queuename=test  2 2`
```

3.12. 集群服务管理页面

本文为您介绍集群的服务及访问地址。

软件	服务	访问地址
Hadoop	yarn resourcemanager	*masternode1_private_ip:8088,masternode2_private_ip:8088*
	jobhistory	*masternode1_private_ip:19888*
	timeline server	*masternode1_private_ip:8188*
	hdfs	*masternode1_private_ip:50070,masternode2_private_ip:50070*
Spark	spark ui	*masternode1_private_ip:4040*
	history	*masternode1_private_ip:18080*
Tez	tez-ui	*masternode1_private_ip:8090/tez-ui2*
Hue	hue	*masternode1_private_ip:8888*
Zeppelin	zeppelin	*masternode1_private_ip:8080*
Hbase	hbase	*masternode1_private_ip:16010*
Presto	presto	*masternode1_private_ip:9090*
Oozie	oozie	*masternode1_private_ip:11000*

软件	服务	访问地址
Ganglia	ganglia	*masternode1_private_ip:9292/ganglia*

3.13. 查看集群列表与详情

本文介绍如何查看您账号下拥有的集群概况和单个集群的详情。

前提条件

已创建集群，详情请参见创建集群。

查看集群列表

1. 登录阿里云E-MapReduce控制台。
2. 在顶部菜单栏处，根据实际情况选择地域和资源组。
3. 单击上方的**集群管理**页签。

   集群列表展示您所拥有的所有集群的基本信息，以及各集群支持的操作。

参数	描述
集群ID/名称	集群的ID以及名称。
集群类型	当前集群的类型。
状态	集群的状态。详情请参见状态表。
创建时间/运行时间	○ 创建时间：显示集群创建的时间。 ○ 运行时间：从创建开始到目前的运行时间。集群一旦被释放，计时终止。
付费类型	集群付费的类型。
集群标签	集群的标识。

参数	描述
操作	集群支持的操作： ○ **详情**：进入集群的详情页面，可以查看集群创建后的详细信息。 ○ **续费**：包年包月集群支持续费。 ○ **更多**： ■ **监控数据**：监控E-MapReduce集群的CPU空闲率、内存容量、磁盘容量等，帮助用户监测集群的运行状态。 ■ **自动续费**：包年包月集群支持自动续费。 ■ **扩容**：集群扩容功能入口，详细信息请参见扩容集群。 ■ **释放**：释放当前集群，仅按量付费集群支持释放，详细信息请参见释放集群。 ■ **重启**：重启当前集群。

查看集群详情

1. 登录阿里云E-MapReduce控制台。
2. 在顶部菜单栏处，根据实际情况选择地域和资源组。
3. 单击上方的**集群管理**页签。
4. 在**集群管理**页面，单击相应集群所在行的**详情**。

   集群基础信息用来展示用户集群的详细信息，包括集群信息、软件信息、网络信息和主机信息四部分。

   ○ 集群信息

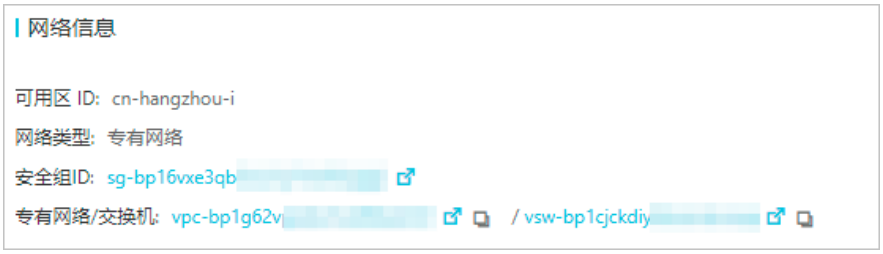
参数	描述
集群名称	集群的名称。
集群ID	集群的实例ID。
地域	集群所在的Region。
当前状态	集群的状态。详情请参见状态表。
IO优化	是否开启了IO优化。
高可用	是否开启了高可用。
安全模式	集群中的软件以Kerberos安全模式启动，详情请参见Kerberos简介。
集群总节点数量	当前集群节点的总数量。
开始时间	集群创建的时间。
付费类型	集群的付费类型。

参数	描述
运行时间	集群运行的时间。
Hive元数据	元数据的方式。
引导操作/软件配置	列出了自定义的脚本和软件的配置信息。
ECS应用角色	列出了ECS应用角色。 当用户程序在EMR计算节点上运行时，可不填写阿里云的AccessKey来访问相关的云服务（例如OSS），EMR会自动申请一个临时AccessKey来授权本次访问。ECS应用角色用于控制这个AccessKey的权限。
标签	集群的标签，详情请参见设置标签。

- 软件信息

软件信息	说明
EMR版本	使用的E-MapReduce版本。
集群类型	当前集群的类型。
软件信息	展示安装的所有应用程序及其版本，例如，HDFS 2.8.5、Hive 2.3.9和Spark 2.4.8。

- 网络信息



网络信息	说明
可用区 ID	集群所在的可用区。
网络类型	默认专有网络。
安全组ID	集群加入的安全组的ID。
专有网络/交换机	当前集群所在的专有网络与子网交换机的ID。

- 主机信息

主机信息	说明
ECS规格	ECS实例规格。
主机数量	当前的节点数量和实际申请的节点数量。理论上这两个值一定是一样的，但是在创建过程中，当前节点会小于申请节点，直到创建完成。
CPU	单个节点的CPU的核数。
内存	单个节点的内存的容量。
数据盘配置	数据盘类型和单个节点的数据盘容量。
ECS实例列表	主实例组包含的ECS实例的信息： ■ **ECS ID**：所购买的ECS的ID。 ■ **组件部署状态**：包含初始化中、正常和扩容中。 ■ **公网**：ECS实例的公网IP地址。 ■ **内网**：ECS实例的内网IP地址，可以被集群中的所有节点访问。 ■ **创建时间**：所购买的ECS的创建时间。

4.服务管理

4.1. 查看服务信息

在集群与服务管理页面，您可以查看集群中的所有服务（例如，HDFS和YARN等）在集群节点上的运行状态。本文为您介绍如果通过控制台查看集群服务信息。

前提条件

已创建集群，详情请参见创建集群。

操作步骤

1. 进入集群管理页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
2. 在**集群管理**页面，单击待查看集群所在行的**集群ID**。
   在**服务列表**区域，即可查看集群中所有服务的运行状态。

> **说明** 服务列表中只会显示创建集群时您勾选的所有服务，未勾选的服务不会显示。

3. 在**服务列表**区域，单击**HDFS**。
   您可以查看对应服务的状态、部署拓扑、配置和配置修改历史。

> **说明** 本文以HDFS服务为例。

4.2. 添加服务

开源大数据平台E-MapReduce支持在控制台增加服务。本文为您介绍如何在E-MapReduce控制台新增服务。

前提条件

已创建集群，详情请参见创建集群。

操作步骤

1. 进入集群与服务管理页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。

iv. 在**集群管理**页面，单击相应集群的**集群ID**。

2. 在**集群与服务管理**页面，单击右上角的**添加服务**。
3. 在**添加服务**对话框，选中待添加的服务，单击**确定**。

   您可以在**服务列表**区域，查看新添加的服务。

4.3. 重启服务

配置项修改后，需要重启对应的服务使配置生效。本文为您介绍如何在E-MapReduce控制台重启服务。

前提条件

已创建集群，详情请参见创建集群。

注意事项

- 为确保服务重启过程中，尽量减少或不影响业务运行，可以通过滚动重启服务。对于有主备状态的实例，会先重启备实例，再重启主实例。
- 滚动执行关闭后，所有节点同时重启可能导致服务不可用，请谨慎选择。

操作步骤

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏，选择**集群服务 > HDFS**。

   本文以HDFS服务为例。
3. 在HDFS服务页面，选择右上角的**操作 > 重启All Components**。
4. 在**执行集群操作**对话框，输入**执行原因**，单击**确定**。
5. 在弹出的**确认**对话框中，单击**确定**。

4.4. 回滚配置

E-MapReduce支持在控制台对已进行过修改的组件参数进行回滚。本文为您介绍如何通过控制台回滚各组件的参数配置。

前提条件

已对组件配置进行过修改操作。

操作步骤

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏，选择**集群服务 > HDFS**。

   本文以HDFS服务为例。

3. 单击上方的**配置修改历史**页签。

   该页面会显示组件的配置历史。

4. 在**配置修改历史**页面，单击待回滚配置操作列的**回滚**。

5. 在**回滚**对话框中，单击**部署生效**。

4.5. 管理组件参数

E-MapReduce提供修改或添加HDFS、YARN和Spark等服务的参数配置功能。本文为您介绍如何修改或添加组件参数。

前提条件

已创建集群，详情请参见创建集群。

修改组件参数

1. 进入详情页面。

   i.

   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。

   iii. 单击上方的**集群管理**页签。

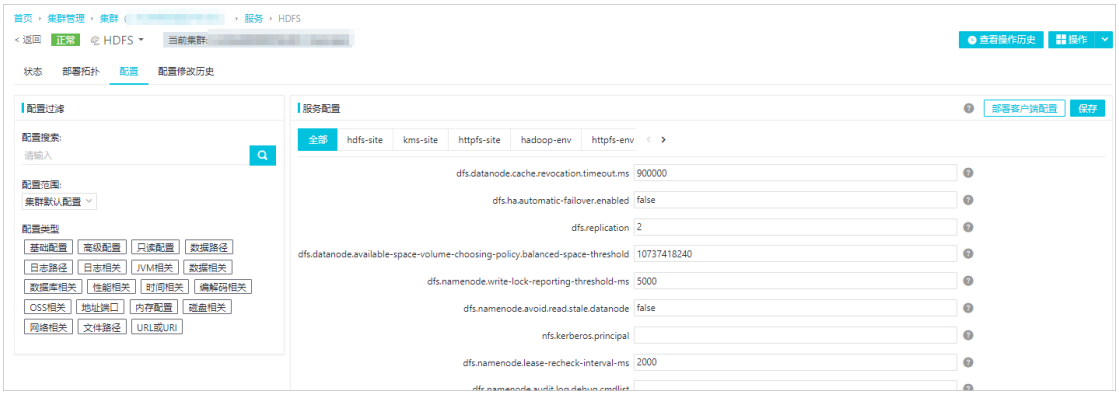
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。

2. 进入**配置**页面。

   i. 单击需要更新参数的组件。

      例如，在左侧导航栏，选择**集群服务 > HDFS**。

ii. 在HDFS服务页面中，单击**配置**页签。



3. 修改配置。

   i. 在**配置搜索**中，输入待修改的配置项，单击 图标。

   ii. 找到您要修改的参数，修改对应的值。

4. 保存配置。

   i. 单击右上角的**保存**。

   ii. 在**确认修改**对话框中，输入执行原因，开启**自动更新配置**。

   iii. 单击**确定**。

5. 生效配置。

   请根据您修改的参数类型执行以下操作，使修改的配置生效。

   - 客户端类型配置

     a. 如果修改的参数类型为客户端类型配置，修改完成后，在**配置**页面，单击**部署客户端配置**。

     b. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。

     c. 在**确认**对话框中，单击**确定**。
     该客户端参数即可生效。 您可以单击上方的**查看操作历史**，查看执行状态和进度。

   - 服务端类型配置

     a. 如果修改的参数类型为服务端类型配置，修改完成后，在**配置**页面，需要单击右上角的**操作**，重启对应的服务。

     b. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。

     c. 在**确认**对话框中，单击**确定**。
     该服务端参数即可生效。 您可以单击上方的**查看操作历史**，查看执行状态和进度。

添加组件参数

1. 进入集群的详情页。

   i.

   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。

   iii. 单击上方的**集群管理**页签。

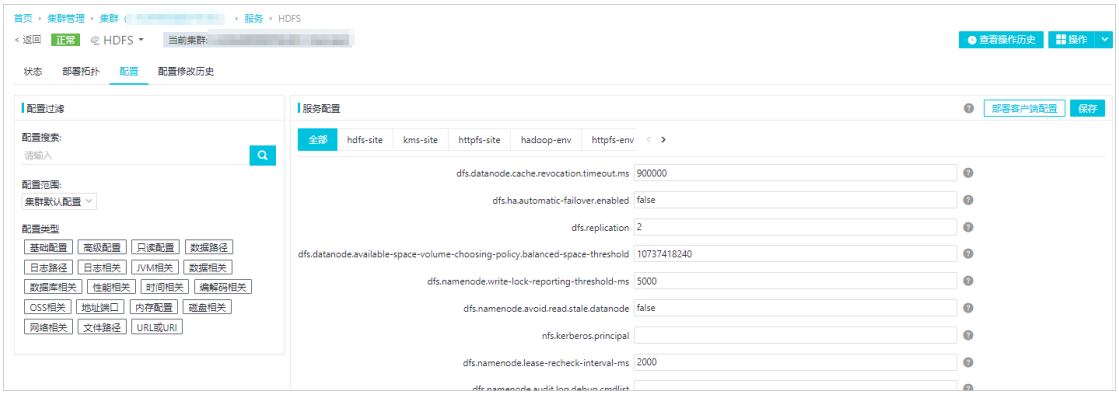
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。

2. 进入**配置**页面。

   i. 单击需要更新参数的组件。

例如，在左侧导航栏选择**集群服务 > HDFS**。

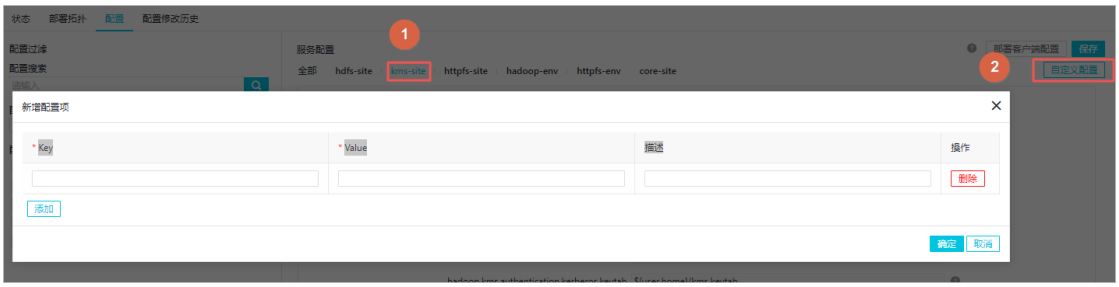
ii. 在HDFS服务页面中，单击**配置**页签。



3. 添加参数配置。

i. 单击上方待添加配置项的页签。

ii. 单击**自定义配置**。



配置项	配置描述
Key	参数名称。
Value	参数值。
描述	参数描述。
操作	支持删除配置项。

iii. 新增完成后，单击**确定**。

4. 保存配置。

i. 单击右上角的**保存**。

ii. 在**确认修改**对话框中，输入执行原因，开启**自动更新配置**。

iii. 单击**确定**。

5. 生效配置。

请根据您修改的参数类型执行以下操作，使修改的配置生效。

- 客户端类型配置

a. 如果添加的参数类型为客户端类型配置，添加完成后，在**配置**页面，单击**部署客户端配置**。

b. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。

c. 在**确认**对话框中，单击**确定**。
该客户端参数即可生效。 您可以单击上方的**查看操作历史**，查看执行状态和进度。

- 服务端类型配置

a. 如果添加的参数类型为服务端类型配置，添加完成后，在**配置**页面，需要单击右上角的**操作**，重启对应的服务。

b. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。

c. 在**确认**对话框中，单击**确定**。
该服务端参数即可生效。 您可以单击上方的**查看操作历史**，查看执行状态和进度。

4.6. 配置自定义软件

Hadoop、Hive和Pig等软件含有大量的配置，当您需要对其软件配置进行修改时，可以在创建集群时通过软件自定义配置功能实现。本文为您介绍如何配置自定义软件。

使用限制

软件配置操作仅在集群创建时执行一次。

操作步骤

1. 进入集群管理页面。
   i.
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
2. 单击右上角的**创建集群**。
3. 在**软件配置**的**高级设置**区域，开启**软件自定义配置**参数。

您可以选择相应的JSON格式配置文件，覆盖或添加集群的默认参数。JSON文件的内容示例如下。

```
[
    {
        "ServiceName":"YARN",
        "FileName":"yarn-site",
        "ConfigKey":"yarn.nodemanager.resource.cpu-vcores",
        "ConfigValue":"8"
    },
    {
        "ServiceName":"YARN",
        "FileName":"yarn-site",
        "ConfigKey":"aaa",
        "ConfigValue":"bbb"
    }
]
```

- ServiceName：服务名，需要全部大写。
- FileName：文件名称，实际传参的文件名称，需要去掉后缀。
- `ConfigKey`：配置项的名称。
- `ConfigValue`：该配置项要设置的具体的值。

各服务的配置文件如下所示。

服务	配置文件
Hadoop	○ core-site.xml ○ log4j.properties ○ hdfs-site.xml ○ mapred-site.xml ○ yarn-site.xml ○ httpsfs-site.xml ○ capacity-scheduler.xml ○ hadoop-env.sh ○ httpfs-env.sh ○ mapred-env.sh ○ yarn-env.sh
Pig	○ pig.properties ○ log4j.properties
Hive	○ hive-env.sh ○ hive-site.xml ○ hive-exec-log4j.properties ○ hive-log4j.properties

集群组件的参数配置好后，您可以继续创建集群，详情请参见创建集群。

4.7. 查看组件部署信息

创建E-MapReduce集群时，不同类型集群的实例节点上会部署不同的服务角色。例如，Hadoop HDFS中的NameNode部署在主实例节点中。本文为您介绍如何查看E-MapReduce集群中服务组件在各节点的部署信息。

前提条件

已创建集群，详情请参见创建集群。

操作步骤

1. 进入详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。

iv. 在**集群管理**页面，单击待查看集群所在行的**详情**。

2. 在左侧导航栏中，单击**集群服务**，选择待查看的集群服务。
3. 在服务页面，单击上方**部署拓扑**页签。

   您可以根据您创建的集群类型，查看服务组件的具体部署信息：

   - Hadoop集群
   - Druid集群
   - Dataflow-Kafka集群
   - Flink集群
   - ZooKeeper集群
   - Data Science集群
   - ClickHouse集群
   - EMR Studio集群

Hadoop集群

以EMR-3.29.0版本为例，Hadoop集群服务组件的具体部署信息如下。

- 必选服务

服务名	主实例节点	核心实例节点
HDFS	○ KMS ○ SecondaryNameNode ○ HttpFS ○ HDFS Client ○ NameNode	○ DataNode ○ HDFS Client
YARN	○ ResourceManager ○ App Timeline Server ○ JobHistory ○ WebAppProxyServer ○ Yarn Client	○ Yarn Client ○ NodeManager
Hive	○ Hive MetaStore ○ HiveServer2 ○ Hive Client	Hive Client
Spark	○ Spark Client ○ SparkHistory ○ ThriftServer	Spark Client
Knox	Knox	无
Tez	○ Tomcat ○ Tez Client	Tez Client

服务名	主实例节点	核心实例节点
Ganglia	○ Gmond ○ Httpd ○ Gmetad ○ Ganglia Client	○ Gmond ○ Ganglia Client
Sqoop	Sqoop Client	Sqoop Client
Bigboot	○ Bigboot Client ○ Bigboot Monitor	○ Bigboot Client ○ Bigboot Monitor
OpenLDAP	OpenLDAP	无
Hue	Hue	无
SmartData	○ Jindo Namespace Service ○ Jindo Storage Service ○ Jindo Client	○ Jindo Storage Service ○ Jindo Client

- 可选服务

服务名	主实例节点	核心实例节点
LIVY	Livy	无
Superset	Superset	无
Flink	○ FlinkHistoryServer ○ Flink Client	Flink Client
RANGER	○ RangerPlugin ○ RangerAdmin ○ RangerUserSync ○ Solr	RangerPlugin
Storm	○ Storm Client ○ UI ○ Nimbus ○ Logviewer	○ Storm Client ○ Supervisor ○ Logviewer
Phoenix	Phoenix Client	Phoenix Client
Kudu	○ Kudu Master ○ Kudu Client	○ Kudu Tserver ○ Kudu Master ○ Kudu Client

服务名	主实例节点	核心实例节点
HBase	○ HMaster ○ HBase Client ○ ThriftServer	○ HBase Client ○ HRegionServer
ZooKeeper	○ ZooKeeper follower ○ ZooKeeper Client	○ ZooKeeper follower ○ ZooKeeper leader ○ ZooKeeper Client
Oozie	Oozie	无
Presto	○ Presto Client ○ PrestoMaster	○ Presto Client ○ PrestoWorker
Impala	○ Impala Runtime and Shell ○ Impala Catalog Server ○ Impala StateStore Server	○ Impala Runtime and Shell ○ Impala Daemon Server
Pig	Pig Client	Pig Client
Zeppelin	Zeppelin	无
FLUME	○ Flume Agent ○ Flume Client	○ Flume Agent ○ Flume Client

Druid集群

以EMR-3.29.0版本为例，Druid集群服务组件的具体部署信息如下。

- 必选服务

服务名	主实例节点	核心实例节点
Druid	○ Druid Client ○ Coordinator ○ Overlord ○ Broker ○ Router	○ MiddleManager ○ Historical ○ Druid Client
HDFS	○ KMS ○ SecondaryNameNode ○ HttpFS ○ HDFS Client ○ NameNode	○ DataNode ○ HDFS Client

服务名	主实例节点	核心实例节点
Ganglia	○ Gmond ○ Httpd ○ Gmetad ○ Ganglia Client	○ Gmond ○ Ganglia Client
ZooKeeper	○ ZooKeeper follower ○ ZooKeeper Client	○ ZooKeeper leader ○ ZooKeeper follower ○ ZooKeeper Client
OpenLDAP	OpenLDAP	无
Bigboot	○ Bigboot Client ○ Bigboot Monitor	○ Bigboot Client ○ Bigboot Monitor
SmartData	○ Jindo Namespace Service ○ Jindo Storage Service ○ Jindo Client	○ Jindo Storage Service ○ Jindo Client

- 可选服务

服务名	主实例节点	核心实例节点
YARN	○ ResourceManager ○ App Timeline Server ○ JobHistory ○ WebAppProxyServer ○ Yarn Client	○ Yarn Client ○ NodeManager
Superset	Superset	无

Dataflow-Kafka集群

以EMR-3.29.0版本为例，Dataflow-Kafka集群服务组件的具体部署信息如下。

- 必选服务

服务名	主实例节点	核心实例节点
Kafka-Manager	Kafka Manager	无
Kafka	○ Kafka Client ○ KafkaMetadataMonitor ○ Kafka Rest Proxy ○ Kafka Broker broker ○ Kafka Schema Registry	○ Kafka Broker broker ○ Kafka Client

服务名	主实例节点	核心实例节点
Ganglia	○ Gmond ○ Httpd ○ Gmetad ○ Ganglia Client	○ Gmond ○ Ganglia Client
ZooKeeper	○ ZooKeeper follower ○ ZooKeeper Client	○ ZooKeeper leader ○ ZooKeeper follower ○ ZooKeeper Client
OpenLDAP	OpenLDAP	无

- 可选服务

服务名	主实例节点	核心实例节点
RANGER	○ RangerPlugin ○ RangerUserSync ○ RangerAdmin ○ Solr	RangerPlugin
Knox	Knox	无

Flink集群

以EMR-3.30.0版本为例，Flink集群服务组件的具体部署信息如下。

- 必选服务

服务名	主实例节点	核心实例节点
HDFS	○ KMS ○ SecondaryNameNode ○ HttpFS ○ HDFS Client ○ NameNode	○ DataNode ○ HDFS Client
YARN	○ ResourceManager ○ App Timeline Server ○ JobHistory ○ WebAppProxyServer ○ Yarn Client	○ Yarn Client ○ NodeManager

服务名	主实例节点	核心实例节点
Ganglia	○ Gmond ○ Httpd ○ Gmetad ○ Ganglia Client	○ Gmond ○ Ganglia Client
ZooKeeper	○ ZooKeeper ○ ZooKeeper Client	○ ZooKeeper ○ ZooKeeper Client
Knox	Knox	无
Flink-Vvp	Flink-Vvp	无
OpenLDAP	OpenLDAP	无

- 可选服务

服务名	主实例节点	核心实例节点
PAI-Alink	Alink	无

ZooKeeper集群

以EMR-3.29.0版本为例，ZooKeeper集群服务组件的具体部署信息如下。

服务名	主实例节点	核心实例节点
Ganglia	无	• Gmond • Httpd • Gmetad • Ganglia Client
ZooKeeper	无	• ZooKeeper leader • ZooKeeper Client • ZooKeeper follower

Data Science集群

以EMR-3.29.1版本为例，Data Science集群服务组件的具体部署信息如下。

- 必选服务

服务名	主实例节点	核心实例节点

服务名	主实例节点	核心实例节点
HDFS	○ HDFS Client ○ KMS ○ HttpFS ○ NameNode ○ SecondaryNameNode	○ HDFS Client ○ DataNode
YARN	○ WebAppProxyServer ○ JobHistory ○ App Timeline Server ○ ResourceManager ○ Yarn Client	○ NodeManager ○ Yarn Client
ZooKeeper	○ ZooKeeper Client ○ ZooKeeper follower	○ ZooKeeper Client ○ ZooKeeper leader ○ ZooKeeper follower
Knox	Knox	无
Tensorflow on YARN	○ TensorFlow-On-YARN-Gateway ○ TensorFlow-On-YARN-History-Server ○ TensorFlow-On-YARN	○ TensorFlow-On-YARN-Client ○ TensorFlow-On-YARN-Gateway
SmartData	○ Jindo Namespace Service ○ Jindo Storage Service ○ Jindo Client	○ Jindo Storage Service ○ Jindo Client
Bigoot	○ Bigboot Monitor ○ Bigboot Client	○ Bigboot Monitor ○ Bigboot Client
PAI-EASYREC	Easyrec	Easyrec
PAI-EAS	PAIEAS	PAIEAS
PAI-Faiss	Faiss	Faiss
PAI-Redis	Redis	Redis
PAI-Alink	Alink	无
Flink-Vvp	Flink-Vvp	无
OpenLDAP	OpenLDAP	无

服务名	主实例节点	核心实例节点
Jindo SDK	Jindo SDK	Jindo SDK

- 可选服务

服务名	主实例节点	核心实例节点
Zeppelin	Zeppelin	无
PAI-REC	Rec	无
AUTOML	AUTOML	AUTOML
TensorFlow	TensorFlow	TensorFlow

ClickHouse集群

以EMR-3.35.0版本为例，ClickHouse集群服务组件的具体部署信息如下。

服务名	主实例节点	核心实例节点
Ganglia	• Gmond • Httpd • Gmetad • Ganglia Client	• Gmond • Ganglia Client
ZooKeeper	• ZooKeeper Client • ZooKeeper follower	• ZooKeeper Client • ZooKeeper leader • ZooKeeper follower
ClickHouse	• ClickHouse Server • ClickHouse Client	• ClickHouse Server • ClickHouse Client

EMR Studio集群

以EMR-3.33.111版本为例，EMR Studio集群服务组件的具体部署信息如下。

服务名	主实例节点	核心实例节点
Ganglia	• Gmetad • Gmond • Httpd • Ganglia Client	• Gmond • Ganglia Client
Zeppelin	Zeppelin Master	Zeppelin Worker
Knox	Knox	
JupyterHub	JupyterHub	

服务名	主实例节点	核心实例节点 无
OpenLDAP	OpenLDAP	
RabbitMQ	RabbitMQ	
Mysql	MySQL	
Airflow	• Airflow Scheduler • Airflow Client • Airflow WebServer	• Airflow Client • Airflow Worker
Studio Workspace	Studio Workspace	无

4.8. 访问Web UI

4.8.1. 通过SSH隧道方式访问开源组件Web UI

本文为您介绍如何通过SSH隧道方式访问开源组件的Web UI。

背景信息

在E-MapReduce集群中，为保证集群安全，Hadoop、Spark和Flink等开源组件的Web UI的端口均未对外开放。您可以通过控制台的方式访问Web UI，也可以通过在本地服务器上建立SSH隧道以端口转发的方式来访问Web UI，端口转发方式包括端口动态转发和本地端口转发两种。

如果您需要通过控制台的方式访问Web UI时，请参见访问链接与端口。

前提条件

- 已创建集群，详情请参见创建集群。
- 确保本地服务器与集群主节点网络连通。您可以在创建集群时打开**挂载公网**开关，或者在集群创建好之后在ECS控制台上为主节点挂载公网，为主节点ECS实例分配固定公网IP或EIP，详情请参见绑定弹性网卡。

获取主节点的公网IP地址

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击目标集群所在行的**详情**。
2. 在**集群基础信息**页面的**主机信息**区域，获取主节点的公网IP地址。

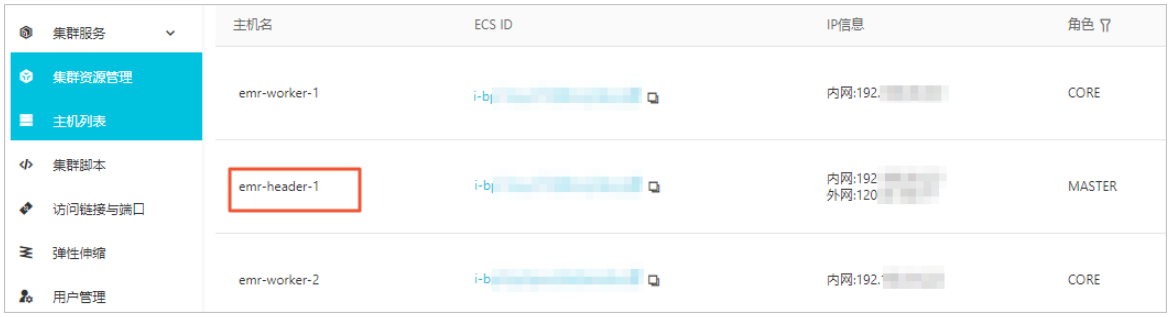
获取主节点的主机名

1. 进入集群详情页面。

i. 登录阿里云E-MapReduce控制台。
ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
iii. 单击上方的**集群管理**页签。
iv. 在**集群管理**页面，单击目标集群所在行的**详情**。

2. 在左侧导航栏，单击**主机列表**。
3. 在**主机列表**页面，查看主节点公网IP地址对应的**主机名**。

主节点IP地址请参见获取主节点的公网IP地址。



使用动态端口转发方式

创建从本地服务器开放端口到集群主节点的SSH隧道，并运行侦听该端口的本地SOCKS代理服务器，端口的数据会由SSH隧道转发到集群主节点。

1. 创建SSH隧道。
   - 密钥方式

     ```
     ssh -i <密钥文件路径> -N -D 8157 root@<主节点公网IP地址>
     ```

   - 密码方式

     ```
     ssh -N -D 8157 root@<主节点公网IP地址>
     ```

     相关参数描述如下：
     - `8157`：本地服务器端口以`8157`为例，实际配置时，您可使用本地服务器未被使用的任意一个端口。
     - `-D`：使用动态端口转发，启动SOCKS代理进程并侦听用户本地端口。
     - `<主节点公网IP地址>`：获取方式请参见获取主节点的公网IP地址。
     - `<密钥文件路径>`：密钥文件保存的路径。
2. 配置浏览器。

   > **注意** 完成隧道创建之后，请保持终端打开状态，此时并不会返回响应。

   完成动态转发配置以后，您可以从以下两种方式中选择一种来进行浏览器配置。
   - Chrome浏览器命令行方式

a. 打开命令行窗口，进入本地Google Chrome浏览器客户端的安装目录。
操作系统不同，Chrome浏览器的默认安装目录不同。

操作系统	Chrome默认安装路径
Mac OS X	*/Applications/Google Chrome.app/Contents/macOS/Google Chrome*
Linux	*/usr/bin/google-chrome*
Windows	*C:\Program Files (x86)\Google\Chrome\Application*

b. 在Chrome浏览器的默认安装目录下，执行以下命令。

```
chrome --proxy-server="socks5://localhost:8157" --host-resolver-rules="MAP * 0.0.0.0 , EXCLUDE localhost" --user-data-dir=/tmp/
```

相关参数描述如下：

- `/tmp/`：如果是Windows操作系统，则`/tmp/`可以写成类似*/c:/tmppath/*的路径。如果是Linux或者Mac OS X操作系统，则可直接写成*/tmp/*路径。
- `8157`：本地服务器端口以`8157`为例，实际配置时，您可使用本地服务器未被使用的任意一个端口。

c. 在浏览器地址栏输入*http://<主节点的主机名>:<port>*，即可访问相应的Web UI。
组件端口信息请参见服务常用端口及配置，主机名的获取请参见获取主节点的主机名。
例如：访问YARN页面时，在浏览器地址栏输入*http://emr-header-1:8088*。

- 代理扩展程序方式
代理扩展程序可以帮助您更加轻松地在浏览器中管理和使用代理，确保网页浏览和集群Web UI访问互不干扰。

a. 安装Chrome的SwitchyOmega插件。

b. 安装完成以后，单击SwitchyOmega插件，然后在弹出框中选择**选项**进行配置。

c. 单击**新建情景模式**，输入**情景模式名称**（例如SSH tunnel），情景模式类型选择**PAC情景模式**。

d. 在**PAC脚本**中配置以下内容。

```
function regExpMatch(url, pattern) {
  try { return new RegExp(pattern).test(url); } catch(ex) { return false; }
}
function FindProxyForURL(url, host) {
    // Important: replace 172.31 below with the proper prefix for your VPC subnet
    if (shExpMatch(url, "*localhost*")) return "SOCKS5 localhost:8157";
    if (shExpMatch(url, "*emr-header*")) return "SOCKS5 localhost:8157";
    if (shExpMatch(url, "*emr-worker*")) return "SOCKS5 localhost:8157";
    return 'DIRECT';
}
```

e. 完成上述参数配置后，在左侧导航栏中单击**应用选项**。

f. 打开Chrome浏览器，在Chrome中单击SwitchyOmega插件，切换到之前创建的SSH tunnel情景模式下。

g. 在浏览器地址栏输入*http://<主节点的主机名>:<port>*，即可访问相应的Web UI。
组件端口信息请参见服务常用端口及配置，主机名的获取方式请参见获取主节点的主机名。
例如：访问YARN页面时，在浏览器地址栏输入*http://emr-header-1:8088*。

使用本地端口转发方式

> **注意** 此方式只能查看最外层的页面，无法查看详细的作业信息。

您可以通过SSH本地端口转发（即将主实例端口转发到本地端口），访问当前主节点上运行的网络应用界面，而不使用SOCKS代理。

1. 在本地服务器终端输入以下命令，创建SSH隧道。
   - 密钥方式

     ```
     ssh -i <密钥文件路径> -N -L 8157:<主节点主机名>:8088 root@<主节点公网IP地址>
     ```

   - 密码方式

     ```
     ssh -N -L 8157:<主节点主机名>:8088 root@<主节点公网IP地址>
     ```

     相关参数描述如下：
     - `-L`：使用本地端口转发，您可以指定一个本地端口，用于将数据转发到主节点本地Web服务器上标识的远程端口。
     - `8088`：主节点上ResourceManager的访问端口，您可以替换为其他组件的端口。
       组件端口信息请参见服务常用端口及配置，主机名的获取请参见获取主节点的主机名。
     - `8157`：本地服务器端口以`8157`为例，实际配置时，您可使用本地服务器未被使用的任意一个端口。
     - `<主节点公网IP地址>`：获取方式请参见获取主节点的公网IP地址。
     - `<密钥文件路径>`：密钥文件保存的路径。
2. 保持终端打开状态，打开浏览器，在浏览器地址栏中输入*http://localhost:8157/*。

服务常用端口及配置

服务	端口	描述
Hadoop 2.X	50070	HDFS Web UI的端口。 配置参数为dfs.namenode.http-address或dfs.http.address。 **说明** dfs.http.address已过期但仍能使用。
	50075	DataNode Web UI的端口。
	50010	Datanode服务端口，用于数据传输。
	50020	IPC服务的端口。
	8020	高可用的HDFS RPC端口。
	8025	ResourceManager端口。 配置参数为yarn.resourcemanager.resource-tracker.address。

Hadoop 2.X 服务	端口	描述
	9000	非高可用的HDFS RPC端口。 配置参数为fs.defaultFS或fs.default.name。 说明 fs.default.name已经过期但仍能使用。
	8088	YARN Web UI的端口。
	8485	JournalNode的RPC端口。
	8019	ZKFC的端口。
	19888	JobHistory Server的Web UI端口。 配置参数为mapreduce.jobhistory.webapp.address。
	10020	JobHistory Server的Web UI端口。 配置参数为mapreduce.jobhistory.address。
Hadoop 3.X	8020	NameNode的端口。 配置参数为dfs.namenode.http-address或dfs.http.address。
	9870	说明 dfs.http.address已过期但仍能使用。
	9871	NameNode的端口。
	9866	DataNode的端口。
	9864	DataNode的端口。
	9865	DataNode的端口。
	8088	ResourceManager的端口。 配置参数为yarn.resourcemanager.webapp.address。
MapReduce	8021	JobTracker的端口。 配置参数为mapreduce.jobtracker.address。
Zookeeper	2181	客户端连接Zookeeper的端口。
	2888	Zookeeper集群内通讯使用，Leader监听此端口。
	3888	Zookeeper端口，用于选举Leader。
	16010	Hbase的Master节点的Web UI端口。 配置参数为hbase.master.info.port。
	16000	HMaster的端口。 配置参数为hbase.master.port。

服务	端口	描述
Hbase	16030	Hbase的RegionServer的Web UI管理端口。 配置参数为hbase.regionserver.info.port。
	16020	HRegionServer的端口。 配置参数为hbase.regionserver.port。
	9099	ThriftServer的端口。
Hive	9083	MetaStore服务默认监听端口。
	10000	Hive的JDBC端口。
	10001	Spark Thrift Sever的JDBC端口。
Spark	7077	• Spark的Master与Worker节点进行通讯的端口。 • Standalone集群提交Application的端口。
	8080	Master节点的Web UI端口，用于资源调度。
	8081	Worker节点的Web UI端口，用于资源调度。
	4040	Driver的Web UI端口，用于任务调度。
	18080	Spark History Server的Web UI 端口。
Kafka	9092	Kafka集群节点之间通信的RPC端口。
Redis	6379	Redis服务端口。
HUE	8888	Hue Web UI的端口。
Oozie	11000	Oozie Web UI的端口。
Druid	18888	Druid Web UI的端口。
	18090	Overlord的端口。 配置参数为**overlord.runtime**页签下的**druid.plaintextPort**。
	18091	MiddleManager的端口。 配置参数为**middleManager.runtime**页签下的**druid.plaintextPort**。
	18081	Coordinator的端口。 配置参数为**coordinator.runtime**页签下的**druid.plaintextPort**。
	18083	Historical的端口。 配置参数为**historical.runtime**页签下的**druid.plaintextPort**。
	18082	Broker的端口。 配置参数为**broker.runtime**页签下的**druid.plaintextPort**。

服务	端口	描述
Ganglia	9292	Ganglia Web UI的端口。
Ranger	6080	Ranger Web UI的端口。
Kafka Manager	8085	Kafka Manager的端口。
Superset	18088	Superset Web UI的端口。
Impala	21050	使用JDBC连接Impala的端口。
Presto	9090	Presto Web UI的端口。

4.8.2. 访问链接与端口

本文介绍如何设置安全组来访问集群上的开源组件。集群创建完成以后，E-MapReduce会为您的集群默认绑定几个域名，方便您访问开源组件的Web UI。

前提条件

已创建E-MapReduce集群，详情请参见创建集群。

> **说明** 确保创建的集群已开启**挂载公网**。

设置安全组访问

当您初次使用组件时，您需要按照以下步骤来打开您的安全组访问权限：

1. 获取机器的公网访问IP地址。

   为了安全的访问集群组件，在设置安全组策略时，推荐您只针对当前的公网访问IP地址开放。获取您当前公网访问IP地址的方法是，访问IP地址，即可查看您当前的公网访问IP地址。

2. 增加安全组策略。

   i. 登录阿里云E-MapReduce控制台。

   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。

   iii. 单击上方的**集群管理**页签。

   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。

   v. 在集群**详情**页面的**网络信息**区域，查看**网络类型**，单击**安全组ID**链接。

vi. 根据需求开启以下端口。

> **注意** 为防止被外部的用户攻击导致安全问题，**授权对象**禁止填写为0.0.0.0/0。

访问不同组件Web UI需要打开的端口如下表。

服务名称	需要打开的端口
YARN UI	8443 **说明** 当集群开启Ranger后，可以访问RANGER UI。
HDFS UI	
Spark History Server UI	
Ganglia UI	
Oozie	
Tez	
ImpalaCatalogd	
ImpalaStatestored	
Storm	
RANGER UI	
Zeppelin	
Hue	8888

以添加**8443**端口为例，执行以下操作：

a. 在**安全组规则**页面，单击**添加安全组规则**。

b. 在**添加安全组规则**对话框中，**端口范围**填写**8443/8443**。

c. **授权对象**填写步骤1中获取的公网访问IP地址。

d. 单击**确定**。

> **说明**
>
> - 如果集群的网络类型是**VPC**，则配置内网入方向。如果集群的网络类型是**经典网络**，则配置公网入方向。此处以**VPC**网络为例。
> - 开放应用出入规则时应遵循最小授权原则。根据应用需求，只开放对应的端口。

vii. 完成以上策略配置后，在**入方向**页面，可查看新增策略。

上述配置成功后，您即已安全的开放了网络访问路径，网络配置完成。

访问开源组件的Web UI

1. 登录阿里云E-MapReduce控制台。
2. 在顶部菜单栏处，根据实际情况选择地域和资源组。
3. 单击上方的**集群管理**页签。
4. 在**集群管理**页面，单击相应集群所在行的**详情**。
5. 在左侧导航栏中，单击**访问链接与端口**。
6. 在**访问链接与端口**页面，单击服务所在行的链接，即可正常的访问Web UI页面。
   - 在2.7.x以上版本，或者是3.5.x以上版本，您可以使用Knox账号访问相应的UI页面，Knox账号创建请参见管理用户，Knox使用请参见Knox 使用说明；访问Hue时请输入Hue的账户和密码，Hue使用请参见使用说明；Zeppelin可直接访问，无需账户密码。
   - 当集群开启Ranger后，您可以使用默认的用户名和密码访问RANGER UI，详细信息请参见概述。
   - 您可以根据集群的版本来访问Flink的Web UI:
     - EMR-3.29.0之前版本
       仅支持通过SSH隧道方式访问Web UI时，请参见通过SSH隧道方式访问开源组件Web UI。

       > **说明** 访问YARN UI页面上的Flink作业：您可以在EMR控制台的**访问链接与端口**页面，单击YARN UI所在行的链接，在Hadoop控制台，单击Flink作业的ID，可以查看Flink作业运行的详情。详情请参见快速入门中的YARN UI方式。

     - EMR-3.29.0及后续版本
       - Flink-Vvp：您可以通过EMR控制台的方式访问Web UI，详情请参见基础使用。
       - Flink（VVR）：您可以通过SSH隧道方式访问Web UI，详情请参见通过SSH隧道方式访问开源组件Web UI。

         > **说明** 访问YARN UI页面上的Flink作业：您可以在EMR控制台的**访问链接与端口**页面，单击YARN UI所在行的链接，在Hadoop控制台，单击Flink作业的ID，可以查看Flink作业运行的详情。详情请参见快速入门中的YARN UI方式。

5.集群运维

5.1. 常用文件路径

本文为您介绍E-MapReduce中常用文件的路径。您可以登录Master节点查看常用文件的安装路径。

大数据组件目录

软件安装目录在*/usr/lib/xxx*下，例如：

- Hadoop：*/usr/lib/hadoop-current*
- Spark：*/usr/lib/spark-current*
- Hive：*/usr/lib/hive-current*
- Flink：*/usr/lib/flink-current*
- Flume：*/usr/lib/flume-current*

您也可以通过登录Master节点，执行**env |grep xxx**命令查看软件的安装目录。

例如，执行以下命令，查看Hadoop的安装目录。

```
env |grep hadoop
```

返回如下信息，其中*/usr/lib/hadoop-current*为Hadoop的安装目录。

```
HADOOP_LOG_DIR=/var/log/hadoop-hdfs
HADOOP_HOME=/usr/lib/hadoop-current
YARN_PID_DIR=/usr/lib/hadoop-current/pids
HADOOP_PID_DIR=/usr/lib/hadoop-current/pids
HADOOP_MAPRED_PID_DIR=/usr/lib/hadoop-current/pids
JAVA_LIBRARY_PATH=/usr/lib/hadoop-current/lib/native:
PATH=/usr/lib/sqoop-current/bin:/usr/lib/spark-current/bin:/usr/lib/hive-current/hcatalog/bin
:/usr/lib/hive-current/bin:/usr/lib/datafactory-current/bin:/usr/local/sbin:/usr/local/bin:/u
sr/sbin:/usr/bin:/usr/lib/b2monitor-current//bin:/usr/lib/b2smartdata-current//bin:/usr/lib/b
2jindosdk-current//bin:/usr/lib/flow-agent-current/bin:/usr/lib/hadoop-current/bin:/usr/lib/h
adoop-current/sbin:/usr/lib/hadoop-current/bin:/usr/lib/hadoop-current/sbin:/root/bin
HADOOP_CLASSPATH=/usr/lib/hadoop-current/lib/*:/usr/lib/tez-current/*:/usr/lib/tez-current/li
b/*:/etc/ecm/tez-conf:/opt/apps/extra-jars/*:/usr/lib/spark-current/yarn/spark-2.4.5-yarn-shu
ffle.jar
HADOOP_CONF_DIR=/etc/ecm/hadoop-conf
YARN_LOG_DIR=/var/log/hadoop-yarn
HADOOP_MAPRED_LOG_DIR=/var/log/hadoop-mapred
```

日志目录

组件日志目录在*/mnt/disk1/log/xxx*下，例如：

- Yarn ResourceManager日志：Master节点*/mnt/disk1/log/hadoop-yarn*
- Yarn NodeNanager日志：Slave节点*/mnt/disk1/log/hadoop-yarn*
- HDFS NameNode日志：Master节点*/mnt/disk1/log/hadoop-hdfs*
- HDFS DataNode日志：Slave节点*/mnt/disk1/log/hadoop-hdfs*
- Hive日志：Master节点*/mnt/disk1/log/hive*
- ESS日志：Master和Worker节点*/mnt/disk1/log/ess/*

配置文件

配置文件目录在*/etc/ecm/xxx*下，例如：

- Hadoop：*/etc/ecm/hadoop-conf/*
- Spark：*/etc/ecm/spark-conf/*
- Hive：*/etc/ecm/hive-conf/*
- Flink：*/etc/ecm/flink-conf/*
- Flume：*/etc/ecm/flume-conf/*

如果您需要修改配置文件中的参数，请登录E-MapReduce控制台操作，通过SSH方式只能浏览配置文件中的参数。

数据目录

JindoFS缓存数据：*/mnt/disk*/bigboot/*

5.2. 登录集群

本文为您介绍如何使用SSH方式（SSH密钥对或SSH密码方式）在Windows和Linux环境中登录集群。

前提条件

- 已创建集群，详情请参见创建集群。
- 确保本地服务器与集群主节点网络连通。您可以在创建集群时打开**挂载公网**开关，或者在集群创建好之后在ECS控制台上为主节点挂载公网，为主节点ECS实例分配固定公网IP或EIP，详情请参见绑定弹性网卡。

背景信息

在本地计算机的终端与集群主节点创建SSH连接之后，您可以通过Linux命令监控集群并与集群交互，也可以在SSH连接中创建隧道以查看开源组件的Web页面，详情请参见通过SSH隧道方式访问开源组件Web UI。

获取主节点公网IP地址

1. 进入详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在**集群基础信息**页面的**主机信息**区域，获取主节点的公网IP地址。

SSH密钥方式

> **说明** 主节点公网IP地址请参见获取主节点公网IP地址。

您可以通过以下三种方式登录集群，详细信息如下：

- 本地使用Linux操作系统

下面步骤以私钥文件*ecs.pem*为例进行介绍：

i. 执行以下命令，修改私钥文件的属性。

```
chmod 400 ~/.ssh/ecs.pem
```

`~/.ssh/ecs.pem` 为*ecs.pem*私钥文件在本地服务器上的存储路径。

ii. 执行以下命令，连接主节点。

```
ssh -i ~/.ssh/ecs.pem root@<主节点公网IP地址>
```

- 本地使用Windows操作系统（通过PuTTY配置信息）
  您可以按照以下方式登录主节点。

i. 下载PuTTY和PuTTYgen。

ii. 将*.pem*私钥文件转换为*.ppk*私钥文件。

a. 运行PuTTYgen。本示例中PuTTYgen版本为0.73。

b. 在**Actions**区域，单击**Load**，导入创建集群时保存的私钥文件。
导入时注意确保导入的格式要求为**All files (*.*)**。

c. 选择待转换的*.pem*私钥文件，单击**打开**。

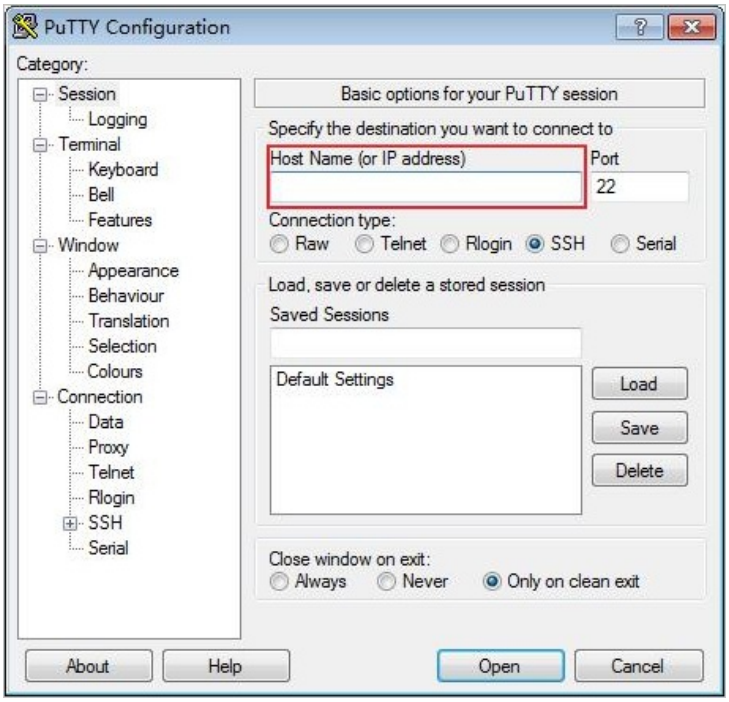
d. 单击**Save private key**。

e. 在弹出的对话框中，单击**是**，指定*.ppk*私钥文件的名称，然后单击**保存**。
保存转化后的私钥到本地。例如：*kp-123.ppk*。

iii. 运行PuTTY。

iv. 选择**Connection > SSH > Auth**，在最下面一个配置项**Private key file for authentication**中，单击**Browse**，选择转化后的密钥文件。

v. 单击**Session**，在**Host Name (or IP address)**下的输入框中，输入登录账号和主节点公网IP地址。
格式为*root@[主节点公网IP地址]*，例如*root@10.10.xx.xx*。



vi. 单击**Open**。

当出现以下提示信息时，说明您已经成功登录实例。

- 本地使用Windows操作系统（通过命令配置信息）
  打开CMD，输入以下命令登录集群。

```
ssh -i <.pem私钥文件在本地机上的存储路径> root@<主节点公网IP地址>
```

SSH密码方式

> **说明** 以下步骤中涉及的用户名，密码分别是root用户和创建集群时设置的密码。主节点公网IP地址请参见获取主节点公网IP地址。

针对不同操作系统，详细的操作步骤如下：

- 本地使用Linux操作系统
  您可以在本地终端的命令行中运行如下命令连接主节点。

```
ssh root@[主节点公网IP地址]
```

- 本地使用Windows操作系统
  i. 下载并安装PuTTY。
  下载链接：PuTTY。
  ii. 启动PuTTY。
  iii. 配置连接Linux实例所需的信息。
    - Host Name (or IP address)：输入实例的固定公网IP或EIP。
    - Port：输入22。
    - Connection Type：选择SSH。
    - （可选）Saved Sessions：输入一个便于识别的名称，然后单击Save即可保存会话，下次登录时无需输入公网IP等信息。
  iv. 单击Open。
  v. 输入用户名（默认为root），然后按回车键。
  输入完成后按回车键即可，登录Linux实例时界面不会显示密码的输入过程。

常见问题

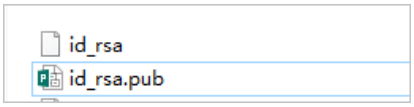
- Q：如何在本地以免密方式登录集群？

A：您可以通过以下步骤在本地以免密方式登录集群。

i. 打开CMD，输入以下命令生成公钥。

```
ssh-keygen
```

本地服务器目录下会生成相应的公钥文件。



ii. 将生成的公钥加入至待访问集群的主节点上。

a. 进入待访问集群的*/.ssh*目录。

```
cd ~/.ssh
```

b. 配置主节点的密钥。

```
vim authorized_keys
```

c. 添加本地公钥*id_rsa.pub*中的内容至*authorized_keys*中。

iii. 加入本地机器的IP地址至安全组。

a. 获取机器的公网访问IP地址。
为了安全的访问集群组件，在设置安全组策略时，推荐您只针对当前的公网访问IP地址开放。获取您当前公网访问IP地址的方法是，访问IP地址，即可查看您当前的公网访问IP地址。

b. 添加安全组规则，以开启22端口。
添加安全组规则，详情请参见添加安全组规则。

iv. 在CMD中，输入以下命令免密登录集群。

```
ssh root@<主节点公网IP地址>
```

- Q：如何登录Core节点？
A：您可以通过以下步骤登录Core节点。

i. 在Master节点上切换到hadoop账号。

```
su hadoop
```

ii. 免密码登录到对应的Core节点。

```
ssh emr-worker-1
```

iii. 通过sudo获得root权限。

```
sudo su - root
```

5.3. 扩容集群

当E-MapReduce集群计算资源或存储资源不足时，您可以对集群进行水平扩展。支持扩展Core节点和Task节点，并且新扩展节点的配置默认与已有节点一致。本文为您介绍如何扩容集群。

背景信息

ClickHouse集群扩容的详情，请参见扩容ClickHouse集群。

前提条件

已创建集群，详情请参见创建集群。

扩容操作

1. 进入集群管理页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
2. 选择目标集群操作列的**更多 > 扩容**。
3. 在**扩容**对话框中，单击**CORE（核心实例组）**或**TASK（任务实例组）**，然后设置相应节点的扩容参数。

以**CORE（核心实例组）**为例，各参数配置项说明如下。

配置项	说明
机器组名称	机器组的名称。
交换机	当前集群的交换机。
配置	当前实例组的配置。
付费类型	当前集群的付费类型。新增节点的付费类型继承集群的付费类型，不可更改。如果是包年包月类型，则您可设置新增节点的**付费时长**。
当前Core数量	默认显示的是当前您所有的Core节点的数量。
增加数量	单击调整框的上下箭头或直接在调整框中输入数字，设置需要增加的Core节点的数量。 **说明** 调整新增Core节点数量过程中，右侧会实时计算集群扩容后的总费用。

配置项	说明
E-MapReduce 服务条款	阅读并同意服务条款后，选中即可。

4. 完成上述参数配置后，单击**扩容**。
5. 查看扩容状态。

   集群扩容操作完成后，在**集群基础信息**页面下方的**主机信息**区域，单击扩容的实例组，在右侧即可查看新增机器的扩容状态。

   当ECS状态为**扩容中**时，说明节点正在扩容中。当ECS状态为**正常**时，说明该节点已加入该集群并可正常提供服务。

修改新增节点密码

集群扩容成功后，您可以通过SSH方式登录新增节点来修改root密码。

1. 登录集群的Master节点，详情请参见登录集群。
2. 执行如下命令，切换到hadoop用户。

   ```
   su hadoop
   ```

3. 执行如下命令，登录新增节点。

   ```
   ssh <ip.of.worker>
   ```

   新增节点的内网IP地址，您可从集群列表与详情页面获取。
4. 执行如下命令，修改root用户密码。

   ```
   sudo passwd root
   ```

5.4. 缩容集群

通过阿里云E-MapReduce（简称EMR）的Task节点缩容功能，可以完成Task节点数量的缩减。

前提条件

已创建集群，详情请参见创建集群。

使用限制

目前仅支持对EMR集群的Task节点缩容，集群还需满足以下条件：

- EMR集群版本2.x高于2.5.0，3.x高于3.2.0。
- 集群状态为空闲或运行中。
- 集群付费类型为按量付费。

操作步骤

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在**集群基础信息**页面，选择右上角的**资源变配 > 缩容**。
3. 在弹出的**缩容**对话框中，单击调整框的下箭头或直接在调整框中输入数字，设置需要保留的Task节点的数量。
4. 完成上述配置后，单击**缩容**。
5. 在**确认缩容机器组**对话框中，单击**确定**。

5.5. 配置弹性伸缩

5.5.1. 弹性伸缩概述

本文介绍E-MapReduce的弹性伸缩功能，您可以根据业务需求和策略设置伸缩策略。该功能目前仅支持E-MapReduce的Hadoop集群，弹性伸缩开启并配置完成后，当业务需求增长时E-MapReduce会自动为您增加Task节点以保证计算能力，当业务需求下降时E-MapReduce会自动减少Task节点以节约成本。

应用场景

在以下场景中，开启E-MapReduce的弹性伸缩功能，可帮助您节省成本，提高执行效率。

- 临时需要按照时间段添加Task节点，补充计算能力。
- 为确保重要作业按时完成，需要按照某些集群指标扩充Task节点。

功能介绍

E-MapReduce支持按时间伸缩和按负载伸缩两种伸缩策略，但使用时两者只能二选一。如果切换伸缩策略，原伸缩规则会保留，但处于失效状态，不会被触发执行；当前已扩容的节点也会保留，除非缩容规则触发，否则不会被缩容。

E-MapReduce弹性伸缩支持抢占式和按量付费两种实例，您可以根据实际需要选择。详细信息请参见管理弹性伸缩。

5.5.2. 新建弹性伸缩机器组

当您的业务量需求不断波动时，建议您开启弹性伸缩功能并配置相应的伸缩规则，以使E-MapReduce可以按业务量波动来增加或减少Task节点。

前提条件

已创建E-MapReduce的Hadoop集群，详细信息请参见创建集群。

使用限制

仅Hadoop集群类型支持弹性伸缩功能。

操作步骤

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏中，选择**弹性伸缩 > 弹性伸缩配置**。
3. 在**弹性伸缩配置**页面，单击**新建弹性伸缩机器组**。
4. 在**新建Task机器组**页面，输入**机器组名称**。

   其他配置可以根据弹性伸缩节点计划的规格来选择，后续配置弹性伸缩的时候也可以适当修改。
5. 选中**E-MapReduce服务条款**复选框。
6. 单击下方的**创建机器组**。

   弹性伸缩配置完成后，后续由于业务波动触发了规则时，您可以在弹性伸缩记录中查看弹性伸缩的历史执行记录以及每次执行的详细结果，详细信息请参见查看弹性伸缩记录。

5.5.3. 管理弹性伸缩

当您的业务量需求不断波动时，建议您开启弹性伸缩功能并配置相应的伸缩规则，以便于E-MapReduce（简称EMR）可以按业务量波动增加或减少Task节点。确保作业完成的同时，可以节省成本。

前提条件

已新建弹性伸缩机器组，创建详情请参见新建弹性伸缩机器组。

配置弹性伸缩

1. 进入详情页面。
   i.
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏中，选择**弹性伸缩 > 弹性伸缩配置**。
3. 在**弹性伸缩配置**页面，单击**配置规则**。
4. 在**弹性伸缩配置**页面，配置如下参数。

i. 在**基础信息**区域，配置相关参数。

参数	描述
最大实例数	弹性伸缩组的Task节点上限。一旦达到上限，即使满足弹性伸缩的规则，也不会继续执行弹性伸缩。最大实例数的上限为500。
最小实例数	弹性伸缩的Task节点下限。如果弹性伸缩规则中设置的增加或减少Task节点数小于此处设置的最小实例数，则在首次执行时，集群会以最小节点数为准进行伸缩。 例如，您设置弹性扩容规则为每天零点动态添加1个节点，最小节点数为3，则系统在第一天的零点时会添加3个节点，以满足最小节点数的要求。
优雅下线	您可以设置超时时间，下线YARN上作业所在的Task节点。如果Task节点没有运行YARN上的作业或者作业运行超出您设置的超时时间，则下线Task节点。超时时间最大值为3600秒。 **说明** 开启优雅下线时，请先将YARN配置项yarn.resourcemanager.nodes.exclude-path的值修改为/etc/ecm/hadoop-conf/yarn-exclude.xml。

ii. 在**成本优化策略**区域，选择成本优化模式。

- **单一计费模式**
  系统会根据您选择的vCPU和内存规格，自动匹配出满足条件的实例，并显示在实例规格区域。您需要选中实例规格区域的实例，以便集群按照已选的实例规格进行伸缩。单一计费模式支持以下两种计费类型：

- **按量付费**
  勾选实例的顺序决定显示节点的优先级，勾选实例后下方区域会显示每一个实例每小时的价格（包含EMR实例价格和ECS实例价格）。

- **抢占式实例**

> **注意** 当您的作业对SLA（Service Level Agreement）要求较高时，请慎用该实例。因为竞价失败等原因会导致抢占式实例被释放，所以请谨慎使用抢占式实例。

勾选实例的顺序决定显示节点的优先级，每一个实例后会显示该实例在按量付费时每小时的价格。您还可以设置每台实例的上限价格，当满足规则时显示该类实例。抢占式实例详情，请参见：抢占式实例概述。

- **成本优化模式**
  该模式下，您可以制定更详细的成本控制策略，在成本和稳定性之间进行调整和权衡。

参数	描述
组内最小按量节点数量	弹性伸缩组需要的按量实例的最小个数，当伸缩组中按量实例个数小于该值时，将优先创建按量实例。
按量节点所占比例	弹性伸缩组内最小按量节点数量满足之后，创建实例中按量实例所占的比例。
最低价的多个实例规格	指定最低价的多个实例规格种类数。当创建抢占实例时，将在这些规格种类中进行均衡分布。最大值为3。
抢占实例补偿	是否开启竞价实例的补偿机制。开启抢占实例补偿后，在竞价实例被回收前5分钟左右，将主动替换掉当前竞价实例。

当您不指定**组内最小按量节点数量**、**按量节点所占比例**和**最低价的多个实例规格**参数时，您创建的是普通成本优化伸缩组。否则，您创建的是成本优化混合实例伸缩组。成本优化混合实例伸缩组与普通成本优化伸缩组在接口和功能方面是完全兼容的。
对于成本优化混合实例伸缩组，您可以通过合理制定混合实例策略，以实现与普通成本优化伸缩组完全相同的行为。例如：

- 普通成本优化伸缩组创建的全为按量实例
  指定**组内最小按量节点数量**=0，**按量节点所占比例**=100，**最低价的多个实例规格**=1。
- 普通成本优化伸缩组优先创建竞价实例
  指定**组内最小按量节点数量**=0，**按量节点所占比例**=0，**最低价的多个实例规格**=1。

> **说明** 成本优化模式详情，请参见AutoScaling成本优化模式升级--混合实例策略。

iii. 在**触发规则**区域，选择规则的触发方式。

- **规定时间伸缩**：请参见按时间伸缩规则配置。
- **规定负载伸缩**：请参见按负载伸缩规则配置。

iv. 单击**确定**。

开启弹性伸缩

弹性伸缩配置完成后，您可以在**弹性伸缩配置**页面，打开**弹性伸缩状态**开关，以开启弹性伸缩。

当弹性伸缩开启之后，如果您修改基础信息或触发规则，在保存配置后，需要在**弹性伸缩配置**页面，单击**应用最新配置**，使修改后的规则生效。

关闭弹性伸缩

> **注意** 当伸缩组内节点数为0时，您才可以关闭弹性伸缩。当伸缩组内节点不为0时，您需要先为伸缩组设置缩容规则或者修改最大实例数为0，直至伸缩组内节点全部缩容，才可以关闭弹性伸缩。

您可以在**弹性伸缩配置**页面，关闭**弹性伸缩状态**开关，以关闭弹性伸缩。

开启弹性伸缩功能后，当您需要更改实例配置或者当您的业务量需求趋于稳定时，您可以关闭弹性伸缩功能。

5.5.4. 按时间伸缩规则配置

如果Hadoop集群计算量在一定的周期内存在明显的波峰和波谷，则您可以设置在每天、每周或每月的固定时间段扩展一定量的Task节点来补充计算能力，这样在保证作业完成的同时，也可以节省成本。

前提条件

已新建弹性伸缩机器组，请参见新建弹性伸缩机器组。

按时间配置伸缩规则

基本信息和成本优化策略的配置详情，请参见管理弹性伸缩。

在E-MapReduce中开启弹性伸缩时，如果选择按时间配置伸缩规则，则根据以下说明配置相关参数即可。

伸缩规则分为扩容规则和缩容规则，本文以扩容规则为例介绍。集群关闭弹性伸缩功能后，所有规则会被清空，再次开启弹性伸缩功能时，需要重新配置伸缩规则。

参数	描述
规则名称	在同一个集群中，伸缩规则名称（包括扩容规则和缩容规则）不允许重复。
规则执行周期	• **重复执行**：您可以选择每天、每周或每月的某一特定时间点执行一次弹性伸缩动作。 • **只执行一次**：集群在指定的时间点执行一次弹性伸缩动作。
重试过期时间(秒)	弹性伸缩在到达指定时间时可能由于各种原因不能执行，通过设置重试过期时间，系统会在该时间范围内每隔30秒尝试执行一次，直到在满足条件时执行伸缩。设置范围为0~21600秒。 假设在指定时间段需要进行弹性伸缩动作A，如果有其他弹性伸缩动作B正在执行或正处在冷却期，则动作A无法执行。在您设置的重试过期时间内，每隔30秒会重试一次，尝试执行A，一旦条件满足，集群会立刻执行弹性伸缩。
扩容(台)	规则被触发时，集群每次增加Task节点数量。
冷却时间(秒)	每次弹性伸缩动作执行完成，到可以再次进行弹性伸缩的时间间隔。在冷却时间内，不会发生弹性伸缩动作。

配置伸缩规格

弹性伸缩配置可以指定伸缩的节点的硬件规格。您只能在开启弹性伸缩功能时配置，保存后不能更改。如果特殊情况确实需要修改，可以关闭弹性伸缩功能后，再次开启。

- 系统会根据您选择的vCPU和内存规格，自动匹配出满足条件的实例，并显示在备选实例列表中。您需要选中备选的实例，以便集群按照已选的实例规格进行伸缩。
- 为避免由于ECS库存不足造成的弹性伸缩失败，您最多可以选择3种ECS实例。
- 无论是选择高效云盘还是SSD云盘，数据盘最小设置为40GB。

5.5.5. 按负载伸缩规则配置

在使用E-MapReduce Hadoop集群时，如果您无法准确的预估大数据计算的波峰和波谷，则可以使用按负载伸缩配置的策略。

前提条件

已新建弹性伸缩机器组，请参见新建弹性伸缩机器组。

按负载配置负载伸缩规则

基本信息和成本优化策略的配置，详情请参见管理弹性伸缩。

伸缩规则分为扩容规则和缩容规则，本文以扩容规则为例介绍。集群关闭弹性伸缩功能后，所有规则会被清空，再次开启弹性伸缩功能时，需要重新配置伸缩规则。切换伸缩策略时（例如，从按负载伸缩切换到按时间伸缩），原策略下的伸缩规则处于失效状态，不会被触发，但已经扩容的节点会继续保留，不会被释放。

参数	描述
规则名称	在同一个集群中，伸缩规则名称（包括扩容规则和缩容规则）不允许重复。
集群负载指标	在YARN的负载指标中获取，详细信息请参见Hadoop官方文档。 E-MapReduce弹性伸缩指标与YARN负载指标的对应关系如E-MapReduce弹性伸缩指标与YARN负载指标的对应关系所示。
统计周期	您选定的集群负载指标在一个统计周期内，按照选定的聚合维度（平均值、最大值和最小值），达到触发阈值为一次触发。
统计规则	

参数	描述
重复几次后扩容	负载指标聚合后达到阈值触发的次数，达到该次数后触发集群弹性伸缩的动作。
扩容(台)	规则被触发时，集群每次执行增加的Task节点数量。
冷却时间(秒)	每次弹性伸缩动作执行完成，到可以再次进行弹性伸缩的时间间隔。在冷却时间内，即使满足弹性伸缩条件也不会发生弹性伸缩动作。即忽略本次在冷却时间内触发的弹性伸缩动作，直到下一次满足伸缩条件且不在冷却时间内再执行。

E-MapReduce弹性伸缩指标与YARN负载指标的对应关系

E-MapReduce弹性伸缩指标	所属服务	说明
YARN.AvailableVCores	YARN	可供分配的虚拟核数。
YARN.PendingVCores	YARN	待分配的虚拟核数。
YARN.AllocatedVCores	YARN	已分配的虚拟核数。
YARN.ReservedVCores	YARN	预留的虚拟核数。
YARN.AvailableMemory	YARN	可供分配的内存量。
YARN.PendingMemory	YARN	待分配的内存量。
YARN.AllocatedMemory	YARN	已分配的内存量。
YARN.ReservedMemory	YARN	预留的内存量。
YARN.AppsRunning	YARN	运行中的任务数。
YARN.AppsPending	YARN	挂起的任务数。
YARN.AppsKilled	YARN	终止的任务数。
YARN.AppsFailed	YARN	失败的任务数。
YARN.AppsCompleted	YARN	完成的任务数。
YARN.AppsSubmitted	YARN	提交的任务数。
YARN.AllocatedContainers	YARN	已分配的容器数。
YARN.PendingContainers	YARN	待分配的容器数。
YARN.ReservedContainers	YARN	预留的容器数。
YARN.MemoryAvailablePrecentage	YARN	剩余内存的百分比 `(MemoryAvailablePrecentage=AvailableMemory/Total Memory)` 。

E-MapReduce弹性伸缩指标	所属服务	说明
YARN.ContainerPendingRatio	YARN	待分配的容器数与已分配的容器数的比率（ContainerPendingRatio = PendingContainers/AllocatedContainers）。

配置伸缩规格

弹性伸缩配置可以指定伸缩的节点的硬件规格。您只能在开启弹性伸缩功能时配置，保存后不能更改。如果特殊情况确实需要修改，可以关闭弹性伸缩功能后，再次开启。

- 系统会根据您选择的vCPU和内存规格，自动匹配出满足条件的实例，并显示在备选实例列表中。您需要选中备选的实例，以便集群按照已选的实例规格进行伸缩。
- 为避免由于ECS库存不足造成的弹性伸缩失败，您最多可以选择3种ECS实例。
- 无论是选择高效云盘还是SSD云盘，数据盘最小设置为40GB。

5.5.6. 查看弹性伸缩记录

本文为您介绍在弹性伸缩执行完成后，如何查看弹性伸缩活动的执行记录。

前提条件

已配置弹性伸缩规则，详细信息请参见创建集群。

操作步骤

1. 进入详情页面。
   i.
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏中，选择**弹性伸缩 > 弹性伸缩记录**。

   您可以查看弹性伸缩活动执行完成后的节点数量和状态等信息。弹性伸缩的状态包括以下五类：
   - **正在执行**：弹性伸缩活动正在执行。
   - **成功**：根据伸缩规则，所有弹性伸缩中的所有节点被加入或移出集群。
   - **部分成功**：根据伸缩规则，有部分节点成功被加入或移出集群，但是受磁盘配额管理或ECS库存的影响，部分节点执行失败。
   - **失败**：根据伸缩规则，没有一个节点被加入或移出集群。
   - **拒绝**：当运行伸缩规则后的实例数大于最大实例数或者小于最小实例数时，或者当运行规则触发时上一次伸缩活动还未结束，就会拒绝该规则运行。

5.5.7. 设置弹性伸缩监控告警

您可以在阿里云的云监控（CloudMonitor）中设置E-MapReduce（简称EMR）弹性伸缩活动的事件告警。当弹性伸缩在运行过程中失败或被拒绝时，云监控系统可以及时地通知组中的联系人，以便及时处理问题。

操作流程

1. 登录云监控控制台。
2. 在左侧导航栏中，单击**事件监控**。
3. 在**事件监控**页面，单击**报警规则**页签。
4. 在**报警规则**页面，单击**创建事件报警**。
5. 在**创建/修改事件报警**对话框中，设置系统事件的报警规则参数。

区域	配置项	描述
基本信息	**报警规则名称**	事件报警规则的名称。
事件报警规则	**产品类型**	选择**E-MapReduce**。
	事件类型	选择**SCALING**。
	事件等级	选择**严重**。
	事件名称	选择全部事件或者具体的某个事件。
	资源范围	事件报警规则作用的资源范围。

其他参数的配置，详细信息请参见创建系统事件报警规则。

6. 单击**确定**。

您可以在**报警规则**页面，查看已创建的规则。

5.6. 扩容磁盘

当E-MapReduce集群的数据存储空间不足时，您可以根据本文进行磁盘（数据盘和系统盘）扩容。本文为您介绍如何对磁盘进行扩容。

背景信息

根据E-MapReduce版本和磁盘属性不同，E-MapReduce支持的磁盘扩容方式也不同，具体说明如下：

- 数据盘：支持在E-MapReduce控制台直接对数据盘进行扩容。不支持缩容。
- 系统盘：支持在ECS控制台对系统盘进行扩容。

数据盘扩容

E-MapReduce控制台仅支持数据盘扩容操作，如果需要对系统盘进行扩容，请参见系统盘扩容。

> **注意** 进行数据盘扩容操作前，请确保当前账号余额充足。数据盘扩容会自动扣款，如果余额不足，则扩容流程会中断。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。

iii. 单击上方的**集群管理**页签。

iv. 在**集群管理**页面，单击相应集群所在行的**详情**。

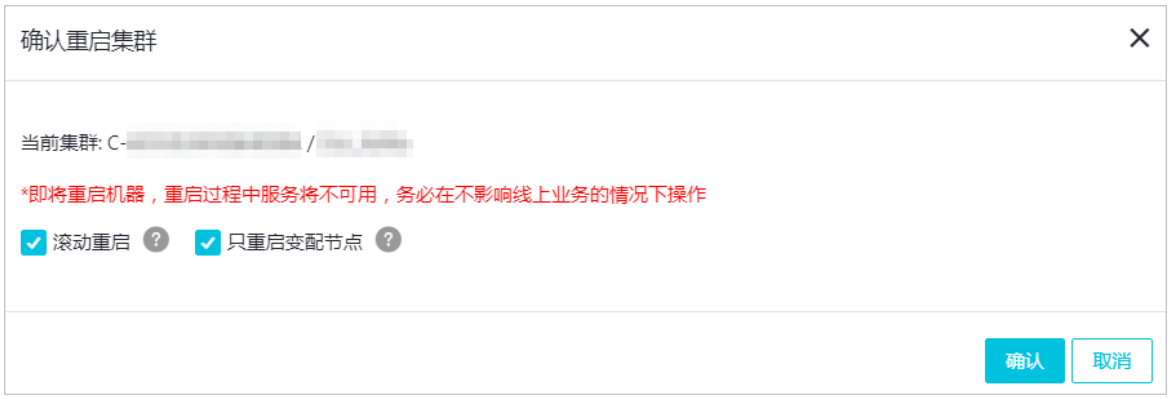
2. 在**集群基础信息**页面的右上方，选择**资源变配 > 磁盘扩容**。
3. 在弹出的**磁盘扩容**对话框中，在**选择机器组**下拉列表选择实例组，修改数据盘容量。

4. 完成配置后，单击**确定**。

磁盘扩容完成后，在**主机信息**区域，实例组会显示**扩容磁盘已完成，重启机器组生效**。

> 注意 磁盘扩容成功后，您需要重启集群以使磁盘扩容生效。重启集群会重启集群中的ECS实例，ECS实例重启过程中大数据服务不可用，请确保在不影响业务的情况下进行重启操作。

5. 单击**扩容磁盘已完成，机器组待重启**，设置集群重启的机制。



集群重启机制配置项。

配置项	说明
滚动重启	**滚动重启**说明如下： ○ 选中**滚动重启**复选框：在一个ECS实例重启完成且该实例上的大数据服务全部恢复后，再启动下一个ECS实例。每个节点重启耗时约5分钟。 ○ 清除**滚动重启**复选框：同时重启ECS实例。
只重启变配节点	变配节点是指已经完成磁盘扩容或者升级配置操作的节点，例如，CORE和MASTER等。**只重启变配节点**说明如下： ○ 选中**只重启变配节点**复选框：只重启变配节点，未变配的节点不会被重启。例如，如果只对**CORE（核心实例组）**的节点做了磁盘扩容，则只重启**CORE（核心实例组）**下的ECS实例。 ○ 清除**只重启变配节点**复选框：重启所有节点，即集群下的所有ECS实例均会重启。

6. 完成上述配置后，单击**确认**，确认重启集群。

> **说明** 扩容完成后，请登录ECS查看磁盘情况，如果扩容没有生效，请重启对应的ECS实例。

系统盘扩容

> **注意** 系统盘扩容是一个比较复杂的操作，如果非特殊需要，建议不要进行系统盘扩容。对于非HA集群，扩容系统盘期间集群不可用。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在**集群基础信息**页面下方的**主机信息**区域，单击待扩容ECS实例的**ECS ID**。
   跳转至ECS控制台。
3. 扩容系统盘。
   i. 在左侧导航栏中，单击**云盘**。
   ii. 在**云盘**页面，**系统盘**所在行的操作列，选择**更多 > 云盘扩容**。
   iii. 在**磁盘扩容**页面，选择**在线扩容**，并设置**扩容后容量**。

   > **说明** 设置的**扩容后容量**不允许小于当前容量。

   iv. 确认费用，阅读并选中**云服务器ECS服务条款**后，单击**确认扩容**。
   v. 在**磁盘扩容须知**对话框中，单击**已阅读，继续扩容**，完成支付。

   在线扩容云盘存在多种限制，如果您的集群不能完全满足这些条件，请使用离线扩容云盘方式。详情请参见在线扩容云盘（Linux系统）或离线扩容云盘（Linux系统）。
4. 系统盘扩容完成后，您需要对扩容的磁盘进行扩展分区和文件系统操作，详情请参见扩展分区和文件系统_Linux系统盘。

   > **说明**
   > - 在扩展分区和文件系统过程中，如果umount命令运行失败，请先在集群上关闭YARN和HDFS服务。
   > - 在Disk1操作时，如果出现ilogtail写日志而无法umount的情况，此时您可以通过sudo pgrep ilogtail | sudo xargs kill命令暂时关闭ilogtail。扩展分区和文件系统完成后，重启节点恢复ilogtail服务。

5. 完成以上操作后，通过SSH方式登录本ECS实例，并通过**df -h**命令查看扩容结果。

> **说明** 系统盘扩容完成后，ECS实例存在以下问题：
>
> - ECS实例会对磁盘做一些处理，这可能导致ECS实例的*/etc/hosts*文件发生变化，您需要在扩容完成后修复*/etc/hosts*文件。
> - SSH免登录配置失效（不影响服务），您可以手动修复。

5.7. 新增机器组

本文为您介绍如何为Task节点新增机器组。

背景信息

您可以新增机器组，以满足不同实例节点的需求。例如，内存型实例节点（vCore：vMem= 1 vCPU：8 GiB）用于大数据离线处理，计算型实例（vCore：vMem=1 vCPU：2 GiB）用于模型训练。

> **注意** 如果您需要在Hadoop集群的Core节点增加机器组，请提交工单处理。

前提条件

已创建Hadoop集群，详情请参见创建集群。

使用限制

E-MapReduce Hadoop集群的Task节点最多支持新增10个机器组。

操作步骤

1. 进入详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在**集群基础信息**页面，选择**资源变配 > 扩容**。
3. 在**扩容**对话框中，单击**TASK（任务实例组）**页签。
4. 在**TASK（任务实例组）**页签，配置以下信息。
   i. （可选）单击**新增机器组**页签。
   ii. 在**新增机器组**页签，输入**机器组名称**。

   > **说明** **机器组名称**在集群中是唯一的，不允许重复。

   iii. 根据您的需求配置各项参数。
   iv. 选中**E-MapReduce服务条款**。
   v. 单击**创建机器组**。

机器组新建完成后，您可以在**集群基础信息**页面的**主机信息**区域查看到。

5.8. 移除异常节点

当组成E-MapReduce集群的ECS节点异常（例如已停止状态），且您不需要该节点时，可移除异常节点。本文为您介绍如何移除异常节点。

前提条件

已创建集群，详情请参见创建集群。

操作步骤

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏中，单击**主机列表**。
3. 找到待移除的实例ID，单击操作列的**删除**。
4. 单击**确定**。
   确认移除实例。

5.9. 升级节点配置

当主实例组或核心实例组的CPU或内存不够时，您可以升级节点配置。本文为您介绍如何升级节点的配置。

前提条件

已创建集群，详情请参见创建集群。

使用限制

- 仅E-MapReduce包年包月集群支持升级配置。

  > **说明** 如果当前是HA集群，只单独重启一台Master，另外一台Master会当做Active的节点来保证服务正常运行。

- 本地盘实例（例如d1和i2实例族）不能升级配置，只能增加节点个数。
- 非本地盘实例只支持升级配置，不支持降低配置。

操作步骤

1. 进入配置升级页面。
   i.
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
   v. 在**集群基础信息**页面，选择**资源变配 > 配置升级**。
2. 修改需要升级的节点配置。
   i. 在**配置升级**页面，根据您的需求修改相应的配置。
   ii. 选中**E-MapReduce服务条款**。
   iii. 单击**确定**。
   等待一段时间会生成订单。
   iv. 支付订单。

在**集群基础信息**页面的**主机信息**区域，会显示如下提示。

3. 在**集群基础信息**页面的**主机信息**区域，单击**升级配置已完成，重启机器组生效**。

> **注意** 因为重启集群会重启集群的ECS实例，所以重启中的ECS实例上的大数据服务不可用，请务必确保不影响业务的情况下操作。

参数	描述
滚动重启	○ 勾选（默认值）：表示一个ECS实例重启完成且该实例上的大数据服务全部恢复后再启动下一个ECS实例。 每个节点重启耗时约五分钟。 ○ 不勾选：表示同时重启所有ECS实例。
只重启变配节点	○ 勾选（默认值）：表示只重启变配节点，未变配的节点不会被重启。 变配节点，是指已经完成过扩容磁盘或者升级配置操作的节点。例如，如果只对核心实例组的节点做了升级配置，并未对主实例组升级配置操作，则只会重启核心实例组下的ECS实例，不会重启主实例组下的ECS实例。 ○ 不勾选：表示所有节点都将重启，即集群下的所有机器都会被重启。

4. 单击**确定**。

重启过程中，在**集群基础信息**页面的**主机信息**区域，升级配置的实例组提示**机器组重启中**。

待提示信息消失后，升级配置全部完成并生效，可以登录集群查验。

> 说明
> - 如果只是升级了CPU而没有升级内存，则升级配置结束。
> - 如果只是升级了内存，或CPU和内存都升级了，建议您执行修改配置，使得YARN可以使用新增的资源。

修改配置

1. 在左侧导航栏中，选择**集群服务 > YARN**。
2. 修改CPU配置。
   i. 在YARN服务页面，单击上方的**配置**页签。
   ii. 在**配置搜索**区域，搜索yarn.nodemanager.resource.cpu-vcores参数并根据您的实际需求修改。
   例如，如果为计算密集型场景，建议调整为ECS vCPU的1：1比例；如果为混合型，可以调到1：2的比例内。如果计算节点为32 vCore且为计算密集型场景，则yarn.nodemanager.resource.cpu-vcores调整为32；如果计算节点为32 vCore且为混合型场景，则yarn.nodemanager.resource.cpu-vcores可以调整到32~64之间。
3. 修改内存配置。
   i. 在YARN服务页面，单击上方的**配置**页签。
   ii. 在**配置搜索**区域，搜索配置项yarn.nodemanager.resource.memory-mb参数，修改配置项的值为 `机器内存*0.8` ，单位为MB。
   例如，新的配置下，内存是32 GB，则需将yarn.nodemanager.resource.memory-mb配置为26214。
4. 保存配置。
   i. 在YARN服务页面，单击右上角的**保存**。
   ii. 在**确认修改**对话框中，输入**执行原因**，单击**确定**。
5. 下发配置。
   i. 在YARN服务页面，选择**操作 > 配置All Components**。
   ii. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。
   iii. 在**确认**对话框中，单击**确定**。
   您可以单击上方的**查看操作历史**，待Configure YARN的任务状态为**成功**之后重启配置。
6. 重启配置。

i. 在YARN服务页面，选择**操作 > 重启All Components**。

ii. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。

iii. 在**确认**对话框中，单击**确定**。

您可以单击上方的单击**查看操作历史**，待Restart YARN的任务状态为**成功**，表示重启配置成功。

5.10. 状态表

本文介绍集群状态表和作业状态列表。

集群状态表

> **说明** 在集群列表和集群详情页面可以看到集群状态。

状态名	状态代码	描述
初始化中	CREATING	集群正在构建，包括两个阶段：一是物理 ECS机器的创建，二是Spark集群的启动，稍等片刻即可达到运行中状态。
创建失败	CREATE_FAILED	创建过程中遇到异常，已经创建的ECS机器会自动回滚，在集群列表页面单击状态右边的问号，可查看异常明细。
运行中	RUNNING	计算集群处于正常运行状态。
空闲	IDLE	集群目前没有运行执行计划。
释放中	RELEASING	单击集群状态列表的释放按钮可达到此状态，此状态表示集群正在努力释放，用户稍等片刻，即可达到已释放状态。
释放失败	RELEASE_FAILED	释放集群过程中出现异常，在集群列表页面单击状态右边的问号，可查看异常明细，出现此状态需要用户重新单击释放按钮。
已释放	RELEASED	计算集群以及托管计算集群的ECS机器均处于释放状态。
状态异常	ABNORMAL	当计算集群中的一个或多个计算节点出现不可恢复的错误时，则集群处于此状态，用户可单击释放集群按钮，释放该集群。
终止中		
终止失败		
已终止		

状态名	状态代码	描述
待支付		
内部状态		
引导操作中		
引导操作失败		
按量集群欠费		

作业状态列表

> **说明** 在作业状态列表里会查看到作业状态。

状态名	状态说明
任务就绪	作业创建信息完整正确，并被成功的保存，正在准备提交到系统的调度队里中，稍等片刻会处于提交中状态。
提交中	作业在计算集群的提交队列中排序等待，还未提交给计算集群进行计算。
提交失败	作业在提交到计算集群的过程中，发生异常，如需再次执行该作业，用户需要克隆作业重新提交。
运行中	作业正在计算集群中拼命的运算，请稍等片刻，单击作业列表中相应的日志按钮可实时查看输出日志。
运行成功	作业在计算集群中被成功的执行并执行完毕，单击作业列表中相应的日志按钮可查看相关日志。
运行失败	作业在计算集群中执行出现异常，单击作业列表中相应的日志按钮可查看相关日志。

5.11. 集群运维指南

本文为您介绍如何查看环境变量和启停E-MapReduce集群服务进程等，以便于您可以自主的运维服务。

前提条件

已创建集群，详情请参见创建集群。

查看环境变量

1. 登录集群，详情请参见登录集群。
2. 输入 `env` 命令。

   您可以看到类似如下的环境变量配置，具体环境变量配置以实际环境为准。

```
PRESTO_HOME=/usr/lib/presto-current
TEZ_CONF_DIR=/etc/ecm/tez-conf
HUDI_HOME=/usr/lib/hudi-current
XDG_SESSION_ID=35918
SPARK_HOME=/usr/lib/spark-current
HOSTNAME=emr-header-1.cluster-23****
HADOOP_LOG_DIR=/var/log/hadoop-hdfs
SMARTDATA_CONF_DIR=/usr/lib/b2smartdata-current//conf
ECM_AGENT_STACK_CACHE_DIR=/usr/lib/emr/ecm-agent/cache/ecm
TERM=xterm
SHELL=/bin/bash
HUE_CONF_DIR=/etc/ecm/hue-conf
HADOOP_HOME=/usr/lib/hadoop-current
FLOW_AGENT_CONF_DIR=/etc/ecm/flow-agent-conf
HISTSIZE=1000
YARN_PID_DIR=/usr/lib/hadoop-current/pids
ECM_AGENT_CACHE_DIR=/usr/lib/emr/ecm-agent/cache
SSH_CLIENT=1.80.115.185 26289 22
HADOOP_PID_DIR=/usr/lib/hadoop-current/pids
EMR_HOME_DIR=/usr/lib/emr
HADOOP_MAPRED_PID_DIR=/usr/lib/hadoop-current/pids
SQOOP_CONF_DIR=/etc/ecm/sqoop-conf
SQOOP_HOME=/usr/lib/sqoop-current
BIGBOOT_MONITOR_HOME=/usr/lib/b2monitor-current/
HCAT_HOME=/usr/lib/hive-current/hcatalog
DATA_FACTORY_CONF_PATH=/etc/ecm/datafactory-conf
HIVE_HOME=/usr/lib/hive-current
PWD=/root
JAVA_HOME=/usr/lib/jvm/java-1.8.0
EMR_DATA_DIR=/usr/lib/emr/data
B2MONITOR_CONF_DIR=/usr/lib/b2monitor-current//conf
HISTCONTROL=ignoredups
SPARK_PID_DIR=/usr/lib/spark-current/pids
SHLVL=1
HOME=/root
HADOOP_MAPRED_LOG_DIR=/var/log/hadoop-mapred
ALLUXIO_CONF_DIR=/etc/ecm/alluxio-conf
ECM_AGENT_LOG_DIR=/usr/lib/emr/ecm-agent/log
TEZ_HOME=/usr/lib/tez-current
DATA_FACTORY_HOME=/usr/lib/datafactory-current
LOGNAME=root
EMR_LOG_DIR=/usr/lib/emr/log
EMR_TMP_DIR=/usr/lib/emr/tmp
XDG_RUNTIME_DIR=/run/user/0
ECM_AGENT_HOME_DIR=/usr/lib/emr/ecm-agent
B2SDK_CONF_DIR=/usr/lib/b2smartdata-current/conf
HIVE_CONF_DIR=/etc/ecm/hive-conf
_=/usr/bin/env
```

登录内置MySQL

1. 通过SSH方式连接集群，详情请参见登录集群。
2. 执行以下命令，登录内置的MySQL。

```
mysql -uroot -pEMRroot1234
```

> **说明** 登录内置MySQL的用户名为root，密码为EMRroot1234。

启停服务进程

您可以在E-MapReduce控制台，对指定服务执行启动、停止和重启操作。各个服务进程的操作类似，下面以HDFS为例，介绍如何启动、停止和重启emr-worker-1主机上的DataNode进程。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域（Region）和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏中，选择**集群服务 > HDFS**。
3. 单击**部署拓扑**页签。
   可以看到该集群中运行的服务进程列表。
4. 操作*emr-worker-1*主机上的DataNode组件。
   i. 启动组件。
      a. 单击**操作**列的**启动**。
      b. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。
      c. 在**确认**对话框中，单击**确定**。
      刷新页面，您可以看到**组件状态**列从STOPPED变为STARTED。
   ii. 重启组件。
      a. 单击**操作**列的**重启**。
      b. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。
      c. 在**确认**对话框中，单击**确定**。
      刷新页面，您可以看到**组件状态**列从STOPPED变为STARTED。
   iii. 停止组件。
      a. 单击**操作**列的**停止**。
      b. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。
      c. 在**确认**对话框中，单击**确定**。
      刷新页面，您可以看到**组件状态**列从STARTED变为STOPPED。

批量操作服务进程

本示例以HDFS为例，介绍如何重启所有实例上的DataNode进程。

1. 进入集群详情页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域（Region）和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击相应集群所在行的**详情**。
2. 在左侧导航栏中，选择**集群服务 > HDFS**。

3. 单击**部署拓扑**页签。
   可以看到该集群中所有运行的服务进程列表。
4. 选择右上角的**操作 > 重启DataNode**。
   i. 在**执行集群操作**对话框中，输入**执行原因**，单击**确定**。
   ii. 在**确认**对话框中，单击**确定**。

> 注意 执行完滚动重启后，不能再执行非滚动重启，否则系统会报错。

通过命令行方式启停服务进程

- YARN
  操作账号：hadoop
  - ResourceManager（Master节点）
    - 启动ResourceManager

      ```
      /usr/lib/hadoop-current/sbin/yarn-daemon.sh start resourcemanager
      ```

    - 停止ResourceManager

      ```
      /usr/lib/hadoop-current/sbin/yarn-daemon.sh stop resourcemanager
      ```

  - NodeManager（Core节点）
    - 启动NodeManager

      ```
      /usr/lib/hadoop-current/sbin/yarn-daemon.sh start nodemanager
      ```

    - 停止NodeManager

      ```
      /usr/lib/hadoop-current/sbin/yarn-daemon.sh stop nodemanager
      ```

  - JobHistoryServer（Master节点）
    - 启动JobHistoryServer

      ```
      /usr/lib/hadoop-current/sbin/mr-jobhistory-daemon.sh start historyserver
      ```

    - 停止JobHistoryServer

      ```
      /usr/lib/hadoop-current/sbin/mr-jobhistory-daemon.sh stop historyserver
      ```

  - WebProxyServer（Master节点）
    - 启动WebProxyServer

      ```
      /usr/lib/hadoop-current/sbin/yarn-daemon.sh start proxyserver
      ```

    - 停止WebProxyServer

      ```
      /usr/lib/hadoop-current/sbin/yarn-daemon.sh stop proxyserver
      ```

- HDFS
  操作账号：hdfs

- NameNode（Master节点）
  - 启动NameNode

    ```
    /usr/lib/hadoop-current/sbin/hadoop-daemon.sh start namenode
    ```

  - 停止NameNode

    ```
    /usr/lib/hadoop-current/sbin/hadoop-daemon.sh stop namenode
    ```

- DataNode（Core节点）
  - 启动DataNode

    ```
    /usr/lib/hadoop-current/sbin/hadoop-daemon.sh start datanode
    ```

  - 停止DataNode

    ```
    /usr/lib/hadoop-current/sbin/hadoop-daemon.sh stop datanode
    ```

- Hive
  操作账号：hadoop
  - MetaStore（Master节点）

    ```
    //启动MetaStore，内存可以根据需要扩大。
    HADOOP_HEAPSIZE=512 /usr/lib/hive-current/bin/hive --service metastore >/var/log/hive/met
    astore.log 2>&1 &
    ```

  - HiveServer2（Master节点）

    ```
    //启动HiveServer2
    HADOOP_HEAPSIZE=512 /usr/lib/hive-current/bin/hive --service hiveserver2 >/var/log/hive/h
    iveserver2.log 2>&1 &
    ```

- HBase
  操作账号：hdfs

  > 注意 需要选择了HBase服务才能使用如下的方式来启动，否则启动的时候会报错。

  - HMaster（Master节点）
    - 启动HMaster

      ```
      /usr/lib/hbase-current/bin/hbase-daemon.sh start master
      ```

    - 重启HMaster

      ```
      /usr/lib/hbase-current/bin/hbase-daemon.sh restart master
      ```

    - 停止HMaster

      ```
      /usr/lib/hbase-current/bin/hbase-daemon.sh stop master
      ```

- HRegionServer（Core节点）
  - 启动HRegionServer

    ```
    /usr/lib/hbase-current/bin/hbase-daemon.sh start regionserver
    ```

  - 重启HRegionServer

    ```
    /usr/lib/hbase-current/bin/hbase-daemon.sh restart regionserver
    ```

  - 停止HRegionServer

    ```
    /usr/lib/hbase-current/bin/hbase-daemon.sh stop regionserver
    ```

- ThriftServer（Master节点）
  - 启动ThriftServer

    ```
    /usr/lib/hbase-current/bin/hbase-daemon.sh start thrift -p 9099 >/var/log/hive/thriftserver.log 2>&1 &
    ```

  - 停止ThriftServer

    ```
    /usr/lib/hbase-current/bin/hbase-daemon.sh stop thrift
    ```

- Hue
  操作账号：hadoop
  - 启动Hue

    ```
    su -l root -c "${HUE_HOME}/build/env/bin/supervisor >/dev/null 2>&1 &"
    ```

  - 停止Hue

    ```
    ps aux | grep hue      //查找到所的Hue进程。
    kill -9 huepid         //终止掉查找到的对应Hue进程。
    ```

- Zeppelin
  操作账号：hadoop
  - 启动Zeppelin

    ```
    //内存可以根据需要扩大。
    su -l root -c "ZEPPELIN_MEM=\"-Xmx512m -Xms512m\" ${ZEPPELIN_HOME}/bin/zeppelin-daemon.sh start"
    ```

  - 停止Zeppelin

    ```
    su -l root -c "${ZEPPELIN_HOME}/bin/zeppelin-daemon.sh stop"
    ```

- Presto
  操作账号：hdfs

- PrestoServer（Master节点）
  - 启动PrestoServer

    ```
    /usr/lib/presto-current/bin/launcher --config=/usr/lib/presto-current/etc/coordinator-config.properties start
    ```

  - 停止PrestoServer

    ```
    /usr/lib/presto-current/bin/launcher --config=/usr/lib/presto-current/etc/coordinator-config.properties stop
    ```

- PrestoServer（Core节点）
  - 启动PrestoServer

    ```
    /usr/lib/presto-current/bin/launcher --config=/usr/lib/presto-current/etc/worker-config.properties start
    ```

  - 停止PrestoServer

    ```
    /usr/lib/presto-current/bin/launcher --config=/usr/lib/presto-current/etc/worker-config.properties stop
    ```

通过命令行方式批量操作

当您需要对Core节点做统一操作时，可以通过脚本命令的方式。在EMR集群中，Master和所有Worker节点在hadoop和hdfs账号下是互通的。

例如，可以通过以下命令，对所有Core节点的NodeManage执行停止操作。本示例Core节点数为10。

```
for i in `seq 1 10`;do ssh emr-worker-$i /usr/lib/hadoop-current/sbin/yarn-daemon.sh stop nodemanager;done
```

5.12. 更换Hadoop集群损坏的本地盘

使用由本地盘机型（i系列和d系列）构建的E-MapReduce（简称EMR）集群时，您可能会收到本地盘受损事件的通知。本文为您介绍如何更换Hadoop集群中损坏的本地盘。

背景信息

如果是emr-worker-1节点或者*/mnt/disk1*盘损坏，请提交工单处理。其他磁盘可以参照本文进行操作。

注意事项

- 整个换盘包括服务停止、卸载磁盘、挂载新盘和服务重启等操作，磁盘的更换通常在五个工作日内完成。执行本文档前请评估服务停止以后，服务的磁盘水位以及集群负载能否承载当前的业务。
- 磁盘更换后，该磁盘上的数据会丢失，请确保磁盘上的数据有足够的副本，并及时备份。

操作步骤

您可以登录ECS控制台，查看事件具体信息，包括实例ID、状态、受损磁盘ID、事件进度和相关的操作。

1. 获取损坏的磁盘信息。
   i. 通过SSH方式连接Hadoop集群，详情请参见登录集群。

ii. 执行以下命令，查看块设备信息。

```
lsblk
```

返回如下类似信息。

```
NAME   MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
vdd    254:48   0  5.4T  0 disk /mnt/disk3
vdb    254:16   0  5.4T  0 disk /mnt/disk1
vde    254:64   0  5.4T  0 disk /mnt/disk4
vdc    254:32   0  5.4T  0 disk /mnt/disk2
vda    254:0    0  120G  0 disk
└─vda1 254:1    0  120G  0 part /
```

iii. 执行以下命令，查看磁盘信息。

```
fdisk -l
```

返回如下类似信息。

```
Disk /dev/vdd: 5905.6 GB, 5905580032000 bytes, 11534336000 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 4096 bytes
I/O size (minimum/optimal): 4096 bytes / 4096 bytes
```

iv. 根据步骤2和步骤3的返回信息记录设备名 `$device_name` 和挂载点 `$mount_path` 。

例如，坏盘事件中的设备为vdd，获取到的设备名为*/dev/vdd*，挂载点为*/mnt/disk3*。

2. 隔离损坏的本地盘。

i. 停止对该坏盘有读写操作的应用。

在EMR管控台停止该节点上会读写本地盘的EMR服务。通常包括HDFS、HBase和Kudu等存储类服务。您也可以在该节点通过以下命令查看占用磁盘的完整进程列表。

```
sudo fuser -mv $device_name
```

ii. 执行以下命令，对本地盘设置应用层读写隔离。

```
sudo chmod 000 $mount_path
```

iii. 执行以下命令，取消挂载本地盘。

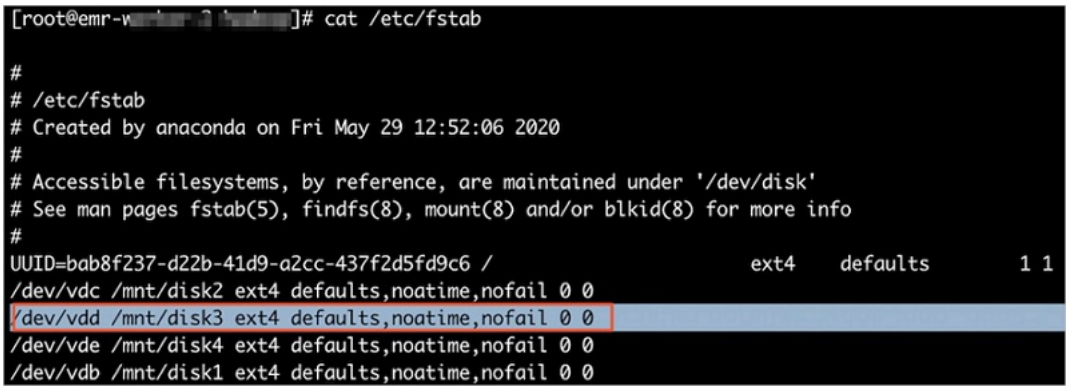
```
sudo umount $device_name;sudo chmod 000 $mount_path
```

> **说明** 如果不执行取消挂载操作，在坏盘维修完成并恢复隔离后，该本地盘的对应设备名会发生变化，可能导致应用读写错误的磁盘。

iv. 更新fstab文件。

a. 备份已有的*/etc/fstab*文件。

b. 删除*/etc/fstab*文件中对应磁盘的记录。
例如，本文示例中坏掉的磁盘是*dev/vdd*，所以需要删除该磁盘对应的记录。



v. 启动步骤2中停止的应用。

3. 换盘操作。

在ECS控制台上修复磁盘，详情请参见隔离损坏的本地盘（控制台）。

4. 挂载磁盘。

磁盘修复完成后，需要重新挂载磁盘，便于使用新磁盘。

i. 执行以下命令，统一设备名。

```
device_name=`echo "$device_name" | sed 's/x//1'`
```

上述命令可以将类似*/dev/xvdk*类的目录名归一化，去掉x，修改为*/dev/vdk*。

ii. 执行以下命令，创建挂载目录。

```
mkdir -p "$mount_path"
```

iii. 执行以下命令，挂载磁盘。

```
mount $device_name $mount_path;sudo chmod 755 $mount_path
```

如果挂载磁盘失败，则可以按照以下步骤操作：

a. 执行以下命令，格式化磁盘。

```
fdisk $device_name << EOF
n
p
1
wq
EOF
```

b. 执行以下命令，重新挂载磁盘。

```
mount $device_name $mount_path;sudo chmod 755 $mount_path
```

iv. 执行以下命令，修改*fstab*文件。

```
echo "$device_name $mount_path $fstype defaults,noatime,nofail 0 0" >> /etc/fstab
```

> **说明** 可以通过 `which mkfs.ext4` 命令，确认ext4是否存在，存在的话 `$fstype` 为ext4，否则 `$fstype` 为ext3。

v. 创建服务目录。

```
mkdir -p $mount_path/data
chown hdfs:hadoop $mount_path/data
chmod 1777 $mount_path/data
mkdir -p $mount_path/hadoop
chown hadoop:hadoop $mount_path/hadoop
chmod 775 $mount_path/hadoop
mkdir -p $mount_path/hdfs
chown hdfs:hadoop $mount_path/hdfs
chmod 755 $mount_path/hdfs
mkdir -p $mount_path/yarn
chown hadoop:hadoop $mount_path/yarn
chmod 755 $mount_path/yarn
mkdir -p $mount_path/kudu/master
chown kudu:hadoop $mount_path/kudu/master
chmod 755 $mount_path/kudu/master
mkdir -p $mount_path/kudu/tserver
chown kudu:hadoop $mount_path/kudu/tserver
chmod 755 $mount_path/kudu/tserver
mkdir -p $mount_path/log
chown hadoop:hadoop $mount_path/log
chmod 775 $mount_path/log
mkdir -p $mount_path/log/hadoop-hdfs
chown hdfs:hadoop $mount_path/log/hadoop-hdfs
chmod 775 $mount_path/log/hadoop-hdfs
mkdir -p $mount_path/log/hadoop-yarn
chown hadoop:hadoop $mount_path/log/hadoop-yarn
chmod 755 $mount_path/log/hadoop-yarn
mkdir -p $mount_path/log/hadoop-mapred
chown hadoop:hadoop $mount_path/log/hadoop-mapred
chmod 755 $mount_path/log/hadoop-mapred
mkdir -p $mount_path/log/kudu
chown kudu:hadoop $mount_path/log/kudu
chmod 755 $mount_path/log/kudu
mkdir -p $mount_path/run
chown hadoop:hadoop $mount_path/run
chmod 777 $mount_path/run
mkdir -p $mount_path/tmp
chown hadoop:hadoop $mount_path/tmp
chmod 777 $mount_path/tmp
```

vi. 使用新磁盘。

在EMR控制台重启在该节点上运行的服务，并检查磁盘是否正常使用。

5.13. 节点下线

当E-MapReduce（简称EMR）集群的节点存在异常，或集群缩容下线节点时，需要在EMR集群管理页面对HDFS DataNode、YARN NodeManager、Jindo Storage Service和HBase HRegionServer服务进行下线（Decommission）操作，如果通过其他方式直接下线节点，可能会导致任务调度失败以及数据安全的问题。本文为您介绍如何进行节点下线操作。

前提条件

已在EMR上创建集群，详情请参见创建集群。

使用限制

当EMR集群的节点存在异常，或集群缩容下线节点时，且您待下线的节点存在HDFS DataNode、YARN NodeManager、Jindo Storage Service或HBase HRegionServer服务，您需要进行节点下线操作。

操作步骤

> **注意** EMR目前支持对HDFS DataNode、YARN NodeManager、Jindo Storage Service和HBase HRegionServer服务进行单节点的下线操作，如果您待下线的节点存在这些服务，则请按照以下顺序进行下线操作。

1. 进入集群管理页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
   iv. 在**集群管理**页面，单击目标集群的集群ID。
2. 下线HDFS DataNode。
   i. 在**集群管理**页面的**服务列表**区域，选择**HDFS**服务所在行的 ... > **Decommission DataNode**。

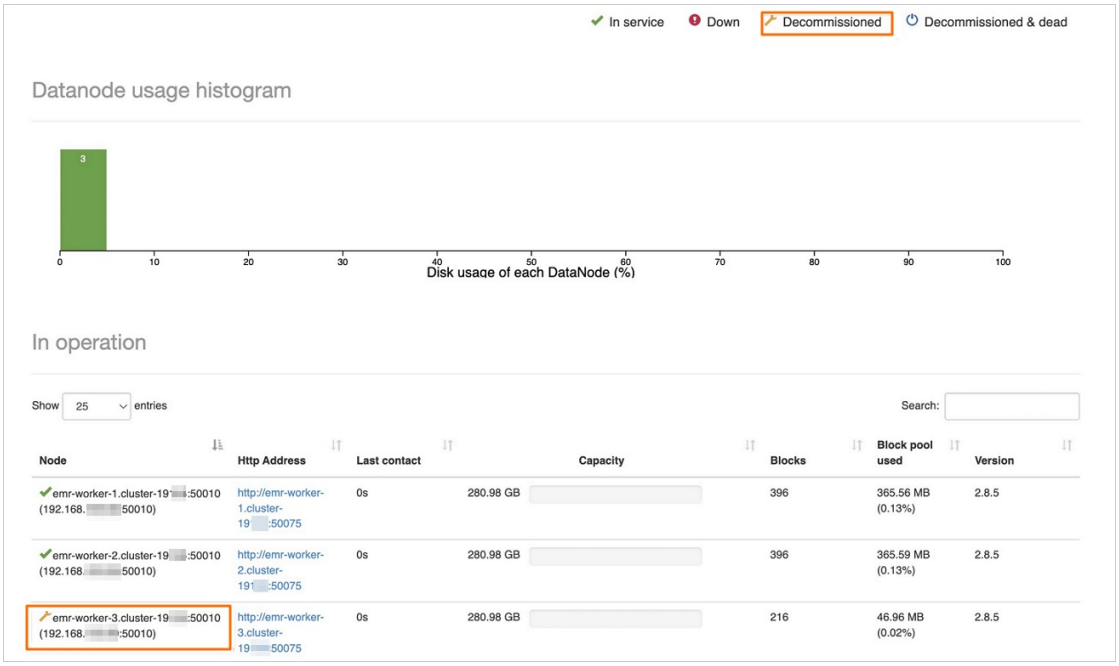
   ii. 在**执行集群操作**对话框中，选择指定机器，输入执行原因，单击**确定**。

> **注意** 确保您待下线的节点存在HDFS DataNode。

   iii. 单击右上角的**查看操作历史**，查看Decommission进度。

iv. 通过HDFS UI查看节点状态。

访问开源组件Web UI的具体操作，请参见访问链接与端口。



3. 下线YARN NodeManager。

i. 在**集群管理**页面的**服务列表**区域，选择YARN服务所在行的 ... > **Decommission DataNode**。

ii. 在**执行集群操作**对话框中，选择指定机器，输入执行原因，单击**确定**。

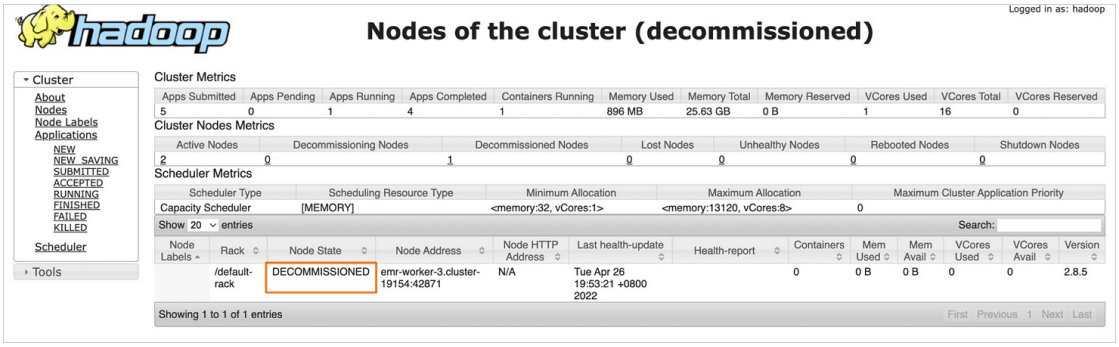
> 注意 确保您待下线的节点存在YARN NodeManager。

iii. 在**确认**对话框中，单击**确定**。

iv. 单击右上角的**查看操作历史**，查看Decommission进度。

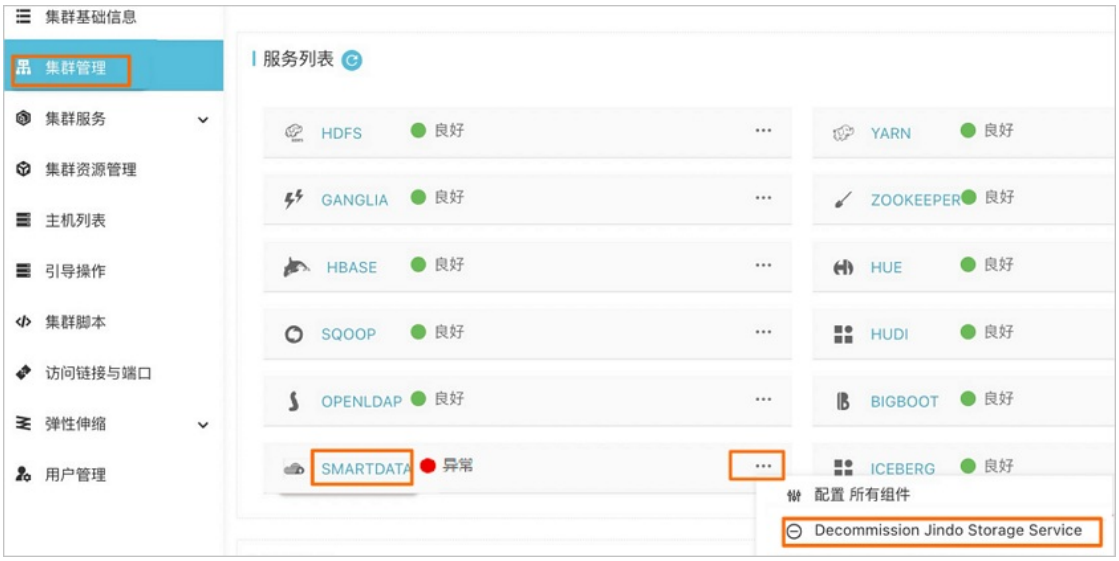
v. 通过YARN UI查看节点状态。

访问开源组件Web UI的具体操作，请参见访问链接与端口。



4. 下线Jindo Storage Service。

i. 在**集群管理**页面的**服务列表**区域，选择SmartData服务所在行的 ... > **Decommission Jindo Storage Service**。



ii. 在**执行集群操作**对话框中，选择指定机器，输入执行原因，单击**确定**。

> 注意 确保您待下线的节点存在Jindo Storage Service。

iii. 在**确认**对话框中，单击**确定**。

iv. 单击右上角的**查看操作历史**，查看Decommission进度。

5. 下线HBase HRegionServer。

i. 在**集群管理**页面的**服务列表**区域，选择**HBase**服务所在行的 ... > **Decommission HRegionServer**。

ii. 在**执行集群操作**对话框中，选择指定机器，输入执行原因，单击**确定**。

> 注意 确保您待下线的节点存在HBase HRegionServer。

iii. 在**确认**对话框中，单击**确定**。

iv. 单击右上角的**查看操作历史**，查看Decommission进度。

5.14. 释放集群

当E-MapReduce按量付费的集群不再使用时，您可以随时释放集群，以节约成本。

前提条件

- 已创建集群，详情请参见创建集群。
- 请确保待释放集群的状态是创建中、运行中或空闲中。

使用限制

仅支持按量付费集群释放。包年包月集群请提交工单处理。

注意事项

确认释放集群后，系统会对集群进行如下处理：

1. 强制终止集群上的所有作业。
2. 终止并释放所有的ECS实例。
   这个过程所需时间取决于集群的大小，集群越小释放越快，一般在几秒内均可完成，至多不会超过5分钟。

操作步骤

1. 进入集群管理页面。
   i. 登录阿里云E-MapReduce控制台。
   ii. 在顶部菜单栏处，根据实际情况选择地域和资源组。
   iii. 单击上方的**集群管理**页签。
2. 在**集群管理**页面，单击待释放集群所在行的**更多 > 释放**。
   您也可以单击待释放集群所在行的**详情**，然后在**集群基础信息**页面，选择**实例状态管理 > 释放**。
3. 在弹出的**集群管理-释放**对话框中，单击**释放**。

6.集群管理常见问题

本文汇总了集群管理时的常见问题。

- 错误提示信息
  - 错误提示：The specified zone is not available for purchase.
  - 错误提示：The request processing has failed due to some unknown error, exception or failure.
  - 错误提示：The Node Controller is temporarily unavailable.
  - 错误提示：Zone或者Cluster的库存不够了
  - 错误提示：指定的InstanceType未授权使用
  - 错误提示：The specified DataDisk Size beyond the permitted range, or the capacity of snapshot exceeds the size limit of the specified disk category.
  - 错误提示：Your account does not have enough balance.
  - 错误提示：The maximum number of Pay-As-You-Go instances is exceeded: create ecs vcpu quota per region limited by user quota [xxx].
  - 错误提示FAILED: SemanticException org.apache.hadoop.hive.ql.metadata.HiveException: java.lang.RuntimeException: Unable to instantiate org.apache.hadoop.hive.ql.metadata.SessionHiveMetaStoreClient
  - 错误提示：User real name authenticate failed.
  - 错误提示：The specified instance Type exceeds the maximum limit for the PostPaid instances.
- 计费问题
  - 为什么集群已续费但还是会收到没有续费的通知？
  - E-MapReduce是否支持自动续费？
  - E-MapReduce如何退款？
- 产品功能咨询问题
  - 集群创建失败需要处理吗？
  - E-MapReduce是否支持竞价实例？
  - 如何申请高配机型？
  - 磁盘容量不足时，该如何处理？
  - 磁盘容量过剩时，该如何处理？
  - 计算能力不足时，该如何处理？
  - 计算能力过剩时，该如何处理？
  - 组件版本过低时，该如何处理？
  - 如何转化非HA集群为HA集群？
  - 如何在EMR上部署第三方软件或服务？
  - 已增加了内存或CPU，为什么YARN上的内存或CPU没有增加？
  - 如何登录Core节点？
  - 集群机器如何分工？
  - 如何处理Kafka集群磁盘异常？
  - E-Mapreduce主节点是否支持安装其他软件？
  - 各个节点上的服务开机会自动启动吗？服务异常中断也会自动重启吗？
  - Hbase的Thrift端口号是多少？

- 之前购买了一套EMR-3.4.3版本，现在需要再购买一套，为什么选不到3.4.3版本？
- E-MapReduce和MaxCompute的区别是什么？
- 存储会自动负载均衡还是需要手动均衡？如果需要手动均衡在哪里操作？
- YARN上ApplicationMaster资源超过Queue限制，该如何处理？
- Worker节点和Header/Gateway节点提交作业的区别是什么？
- 如何访问开源组件的Web UI?

- 元数据管理问题
  - Metastore初始化时提示Failed to get schema version异常信息，该如何处理？
  - 如果Hive元数据信息中包含中文信息，例如列注释和分区名等，该如何处理？
- 权限问题
  - 经典网络中ECS与VPC中的E-MapReduce集群如何进行网络互访？
  - 同账号不同VPC下的E-MapReduce如何互连？
  - 如何对不同RAM用户的OSS数据进行隔离？

错误提示：The specified zone is not available for purchase.

您选择创建集群的可用区暂时停止了按量计费ECS的售卖，建议您更换可用区购买。

错误提示：The request processing has failed due to some unknown error, exception or failure.

您可以稍等一会重试，也可以立即提交工单处理。

错误提示：The Node Controller is temporarily unavailable.

请稍等一会再重试创建集群。

错误提示：Zone或者Cluster的库存不够了

待创建E-MapReduce集群可用区的ECS实例库存不足。您可以尝试手动选择其他可用区重新创建集群，或者使用随机模式创建集群。

错误提示：指定的InstanceType未授权使用

ECS的按量付费高配机型（8核以上的所有机型）需要用户申请开通以后才可以使用，请单击提交工单申请。请申请8核16 GiB、8核32 GiB、16核32 GiB和16核64 GiB这四种目前E-MapReduce使用的机型。

错误提示：The specified instance Type exceeds the maximum limit for the PostPaid instances.

问题分析：

- 您的按量节点数量已达到上限。
- 您没有创建这个机型的权限。

解决方法：

- 申请增加节点数量。
- 在ECS开通这个机型的使用权限。

如何登录Core节点？

1. 在Master节点上切换到hadoop账号。

```
su hadoop
```

2. 免密码登录到对应的Core节点。

```
ssh emr-worker-1
```

3. 通过sudo获得root权限。

```
sudo su - root
```

为什么集群已续费但还是会收到没有续费的通知？

问题分析：因为E-MapReduce涉及E-MapReduce和ECS两部分费用，大部分的用户都只是续费了ECS而没有续费E-MapReduce。

解决方法：您可以进入**集群基础信息**页面，在**费用管理 > 续费**页面，查看ECS和E-MapReduce到期时间。

E-MapReduce是否支持自动续费？

E-MapReduce支持自动续费操作，包括E-MapReduce和ECS的自动续费。

E-MapReduce如何退款？

如果您需要申请退款，请提交工单处理。

集群创建失败需要处理吗？

可以不用处理。对应的计算资源并没有真正地创建出来。创建失败的集群三天后自动隐藏。

集群机器如何分工？

E-MapReduce中包含一个Master节点和多个Slave（或者Worker）节点。其中Master节点不参与数据存储和计算任务，Slave节点用来存储数据和计算任务。例如，3台4核8 GiB机型的集群，其中一台机器用来作为Master节点，另外两台用来作为Slave节点，也就是集群的可用计算资源为2台4核8 GiB机型。

如何处理Kafka集群磁盘异常？

问题分析：常见磁盘异常包括磁盘写满和磁盘损坏。

解决方法：

- 磁盘写满：
  i. 登录服务器。
  ii. 查找到写满的磁盘，然后按照以下原则清理数据。
    - 切勿直接删除Kafka的数据目录，否则所有数据会丢失。
    - 切勿清理Kafka的系统topic，例如consumer_offsets和schema等。
    - 查找到占用空间较多或者明确不需要的topic，选择其中某些Partition，从最早的日志数据开始删除。删除Segment及相应的index和timeindex文件。
  iii. 重启Kafka broker服务。
- 磁盘损坏：
  - 磁盘损坏不超过25%时，无需处理。
  - 磁盘损坏超过25%时，请提交工单处理。

错误提示：The specified DataDisk Size beyond the permitted range, or the capacity of snapshot exceeds the size limit of the specified disk category.

磁盘参数设置太小，建议您调整到40 GB以上重试。

错误提示：Your account does not have enough balance.

账号余额不足。

错误提示：The maximum number of Pay-As-You-Go instances is exceeded: create ecs vcpu quota per region limited by user quota [xxx].

您按量计费的ECS使用量达到上限，需要申请提高ECS使用量上限，或者释放部分按量计费的ECS以便继续创建E-MapReduce集群。

E-Mapreduce主节点是否支持安装其他软件？

不建议安装。因为安装其他软件可能会影响集群的稳定性和可靠性，所以不建议安装。

各个节点上的服务开机会自动启动吗？服务异常中断也会自动重启吗？

会自动启动。服务异常中断会自动恢复。

Hbase的Thrift端口号是多少？

Hbase的Thrift端口号是9099。

E-MapReduce是否支持竞价实例？

E-MapReduce在使用弹性伸缩功能时，支持抢占式实例，详情请参见管理弹性伸缩。

错误提示 `FAILED: SemanticException org.apache.hadoop.hive.ql.metadata.HiveException: java.lang.RuntimeException: Unable to instantiate org.apache.hadoop.hive.ql.metadata.SessionHiveMetaStoreClient`

问题描述：创建集群时选择了**统一meta库**，但执行Hive报错。

通常是因为集群没有公网EIP，导致集群未创建成功。您可以提交工单处理。

错误提示：User real name authenticate failed.

用户账号没有完成实名认证。请到阿里云上完成即可。E-MapReduce部分也会有这个实名认证，可以提交工单处理。

之前购买了一套EMR-3.4.3版本，现在需要再购买一套，为什么选不到3.4.3版本？

E-MapReduce的主版本会定期升级，对于部分老版本会做下线处理。您可以查看现有E-MapReduce版本中各个软件的版本（例如Hive和Spark）是否满足您的需求。如果您需要已下线的EMR版本，请提交工单处理。

E-MapReduce和MaxCompute的区别是什么？

两者都是大数据处理解决方案。E-MapReduce是完全在开源基础上构建的大数据平台，对开源软件的使用方式和实践方式100%兼容。MaxCompute由阿里巴巴开发，对外不开源，但是封装后使用起来比较方便，而且运维成本也较低。

存储会自动负载均衡还是需要手动均衡？如果需要手动均衡在哪里操作？

需要手动均衡，操作步骤如下：

1. 登录阿里云E-MapReduce控制台。
2. 在顶部菜单栏处，根据实际情况选择地域和资源组。
3. 单击上方的**集群管理**页签。
4. 在**集群管理**页面，单击相应集群所在行的**详情**。
5. 单击左侧导航栏的**集群服务 > HDFS**。
6. 单击右上角的**操作 > Rebalance**。

7. 在**执行集群操作**对话框中，配置各项参数，单击**确定**。
8. 在**确认**对话框中，单击**确定**。

如何申请高配机型？

请提交工单处理。

磁盘容量不足时，该如何处理？

因为EMR集群不支持磁盘数量的增加，所以您可以在EMR控制台调大单块磁盘的容量，或扩容Core节点。

磁盘容量过剩时，该如何处理？

因为EMR集群不支持磁盘容量缩容，所以您可以重新购买集群，详情请参见创建集群。

计算能力不足时，该如何处理？

您可以在EMR控制台上扩容Task节点，详情请参见扩容集群。

计算能力过剩时，该如何处理？

根据集群区分如下：

- 针对按量付费集群，您可以在EMR控制上缩容Task节点。
- 针对包年包月集群，先在Task节点decommission YARN Nodemanager服务，然后在ECS控制台将对应Task节点的ECS转为按量付费后再释放。

组件版本过低时，该如何处理？

因为EMR集群不支持单组件版本升级，所以您可以重新购买高版本的集群，详情请参见创建集群。

如何转化非HA集群为HA集群？

EMR集群暂不支持非HA集群转为HA集群，建议您重新购买HA集群。

如何在EMR上部署第三方软件或服务？

建议您在集群创建时通过引导操作安装第三方软件或服务。如果集群创建后手工安装第三方软件或服务，在扩容时，新扩容节点需重新手工安装第三方软件或服务。

当集群的ECS发出告警，提示错误需要重新部署时如何处理？

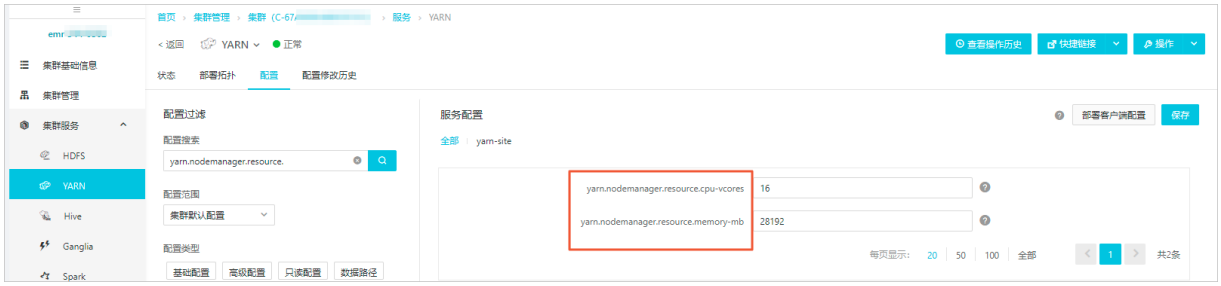
请提交工单处理。

已增加了内存或CPU，为什么YARN上的内存或CPU没有增加？

因为YARN上的资源是由每个NodeManager的资源相加得到的总值，每个NodeManager的值是固定的，所以您可以通过修改配置并重启NodeManager，来影响YARN的总体资源，修改组件参数详情请参见管理组件参数。

修改配置如下：

- 内存：yarn.nodemanager.resource.memory-mb
- CPU：yarn.nodemanager.resource.cpu-vcores



YARN上ApplicationMaster资源超过Queue限制，该如何处理？

- 问题现象：
  - YARN上任务一直显示accepted。
  - 错误提示 `Application is added to the scheduler and is not yet activated. Queue's AM resource limit exceeded.` 。
- 处理方法：任务在YARN上运行时会先启动一个ApplicationMaster，由这个Master来申请运行数据的资源，但是有一个参数会限制Master占用YARN资源的百分比，默认是0.25（25%）。如果您的作业都是消耗资源比较小的任务，建议可以修改yarn.scheduler.capacity.maximum-am-resource-percent的值，调整到0.5~0.8之间，修改参数后无需重启服务，您可以直接在节点上执行 `yarn rmadmin -refreshQueues` 命令。修改组件参数详情请参见管理组件参数。

Worker节点和Header/Gateway节点提交作业的区别是什么？

在数据开发的作业编辑页面，单击右上角的**作业设置**，在**作业设置**页面的**高级设置**页签，您可以在**模式**区域，选择**在Worker节点提交**或者**在Header/Gateway节点提交**。

在Worker节点提交：不可以指定机器提交，会先启动一个名为Launcher的任务，再来启动和监控您实际的任务，即需要用到两个ApplicationMaster的资源，通常这两个任务的ApplicationId是连续的。

在Header/Gateway节点提交：可以指定机器提交，仅启动实际的任务，但是如果任务太多，会对Header/Gateway节点造成一定的压力，如果机器内存不足，任务无法启动。

经典网络中ECS与VPC中的E-MapReduce集群如何进行网络互访？

目前阿里云存在经典网络和VPC两种网络类型。由于E-MapReduce集群是在VPC网络中，而很多用户的业务系统还存在于经典网络中，为了解决此问题，阿里云推出了ClassicLink方案，您可以参见此方案进行网络互访，详情请参见ClassicLink方案。

ClassicLink方案的简要操作步骤如下：

1. 指定网段创建vSwitch，详情请参见ClassicLink方案。
2. 创建集群时，使用指定网段的vSwitch来部署E-MapReduce集群。
3. 在ECS控制台，将对应的经典网络节点连接到这个VPC。
4. 设置安全组访问规则。

如何访问开源组件的Web UI？

您可以通过E-MapReduce控台的**访问链接与端口**功能，访问开源组件的Web UI，详情请参见访问链接与端口。

如何对不同RAM用户的OSS数据进行隔离？

您可以使用访问控制RAM（Resource Access Management），对不同RAM用户的OSS数据进行隔离。操作步骤如下：

1. 通过阿里云账号登录RAM控制台。
2. 创建RAM用户。
   i. 在左侧导航栏中，选择**身份管理 > 用户**。
   ii. 单击**创建用户**。

   > **说明** 可以一次性创建多个RAM用户。

   iii. 输入**登录名称**和**显示名称**。
   iv. 在**访问方式**区域，选择访问方式。
   - **控制台访问**：完成对登录安全的基本设置，包括自动生成或自定义登录密码、是否要求下次登录时重置密码以及是否要求开启多因素认证。
   - **OpenAPI调用访问**：自动为RAM用户生成访问密钥（AccessKey），支持通过API或其他开发工具访问阿里云。

   > **说明** 为了保障账号安全，建议仅为RAM用户选择一种登录方式，避免RAM用户离开组织后仍可以通过访问密钥访问阿里云资源。

   v. 单击**确定**。
3. 新建权限策略。
   i. 在左侧导航栏中，选择**权限管理 > 权限策略管理**。
   ii. 单击**创建权限策略**。
   iii. 填写**策略名称**。

iv. 选中**脚本配置**。**脚本配置**方法请参见语法结构编辑策略内容。 本示例分别按照不同环境来创建两个权限策略，您可以根据自己的环境选择对应的脚本进行创建。

测试环境（test-bucket）	生产环境（prod-bucket）

<pre>{
"Version": "1",
"Statement": [
{
"Effect": "Allow",
"Action": [
"oss:ListBuckets"
],
"Resource": [
"acs:oss:*:*:*"
]
},
{
"Effect": "Allow",
"Action": [
"oss:Listobjects",
"oss:GetObject",
"oss:PutObject",
"oss:DeleteObject"
],
"Resource": [
"acs:oss:*:*:test-bucket",
"acs:oss:*:*:test-bucket/*"
]
}
]
}</pre>
	<pre>{
	"Version": "1",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": [
	"oss:ListBuckets"
	],
	"Resource": [
	"acs:oss:*:*:*"
	]
	},
	{
	"Effect": "Allow",
	"Action": [
	"oss:Listobjects",
	"oss:GetObject",
	"oss:PutObject"
	],
	"Resource": [
	"acs:oss:*:*:prod-bucket",
	"acs:oss:*:*:prod-bucket/*"
	]
	}
	]
	}</pre>

按上述脚本示例进行权限隔离后，RAM用户在E-MapReduce控制台的限制如下：

- 在创建集群、创建作业和创建工作流的OSS文件页面，可以看到所有的Bucket，但是只能进入被授权的Bucket。
- 只能看到被授权的Bucket下的内容，无法看到其他Bucket内的内容。
- 作业中只能读写被授权的Bucket，读写未被授权的Bucket会报错。

v. 单击**确定**。

4. （可选）为RAM用户授权。
创建的RAM用户未授权，则您可按以下方法进行授权：
   i. 在左侧导航栏，选择**身份管理 > 用户**。
   ii. 单击待授权RAM用户所在行的**添加权限**。
   iii. 单击需要授予RAM用户的权限策略，单击**确定**。
   iv. 单击**完成**。
5. （可选）开启RAM用户的控制台登录权限。
如果创建RAM用户时未开启控制台登录权限，则您可按以下方法进行开启：
   i. 单击左侧导航栏，选择**身份管理 > 用户**。

ii. 单击目标RAM用户的用户登录名称。

iii. 在**控制台登录管理**区域，单击**修改登录设置**。

iv. 在**修改登录设置**页签，开启**控制台访问**。

v. 单击**确定**。

6. 通过RAM用户登录E-MapReduce控制台。

i. 通过RAM用户登录控制台。

ii. 登录控制台后，选择**E-MapReduce**产品即可。

同账号不同VPC下的E-MapReduce如何互连？

专有网络VPC（Virtual Private Cloud）可以为您创建一个隔离的网络环境，支持自定义IP地址范围、划分网络、配置路由表和网关等。您可以将多个E-MapReduce集群创建在不同的VPC下，通过高速通道的配置使其可以互连。

创建专有网络集群相关信息如下：

- VPC：选择当前创建的E-MapReduce集群归属的VPC，如果还没创建可以进入VPC控制台进行创建，通常一个账号最多创建2个VPC网络，超过2个时需提交工单处理。
- 交换机：E-MapReduce集群内的ECS实例通过交换机进行通信，如果还没创建可以进入VPC控制台，单击**交换机**页签进行创建。因为交换机有可用区的属性，所以创建交换机的可用区需要和在E-MapReduce创建集群时一致。
- 安全组名称：集群所属的安全组，经典网络的安全组不能在VPC中使用，VPC的安全组只能在当前VPC中使用。安全组列表中只展示您在E-MapReduce产品中创建的安全组。如果需要新建安全组，直接输入安全组名称即可。

以下通过示例为您介绍，创建两个不同VPC下的E-MapReduce集群，并通过云企业网配置使其中一个集群可以访问另一个集群（Hive访问HBase）。

1. 创建集群，详情请参见创建集群。
   在E-MapReduce控制台上创建两个集群，Hive集群C1处于VPC1中，HBase集群C2处于VPC2中，两个集群都在杭州区域。
2. 配置同账号VPC互连。
   详细配置信息，请参见创建云企业网实例。
3. 使用SSH登录HBase集群，通过HBase Shell创建表。

```
create 'testfromHbase','cf'
```

4. 使用SSH登录Hive集群，进行以下配置，即可实现同账号不同VPC下的E-MapReduce集群互连。

i. 修改hosts，增加如下信息。

```
$zk_ip emr-cluster //$zk_ip为HBase集群的ZooKeeper节点IP地址。
```

ii. 通过Hive Shell访问HBase。

```
set hbase.zookeeper.quorum=172.*.*.111,172.*.*.112,172.*.*.113;
CREATE EXTERNAL TABLE IF NOT EXISTS testfromHive (rowkey STRING, pageviews Int, bytes
STRING) STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler' WITH SERDEPROPER
TIES ('hbase.columns.mapping' = ':key,cf:c1,cf:c2') TBLPROPERTIES ('hbase.table.name'
= 'testfromHbase');
```

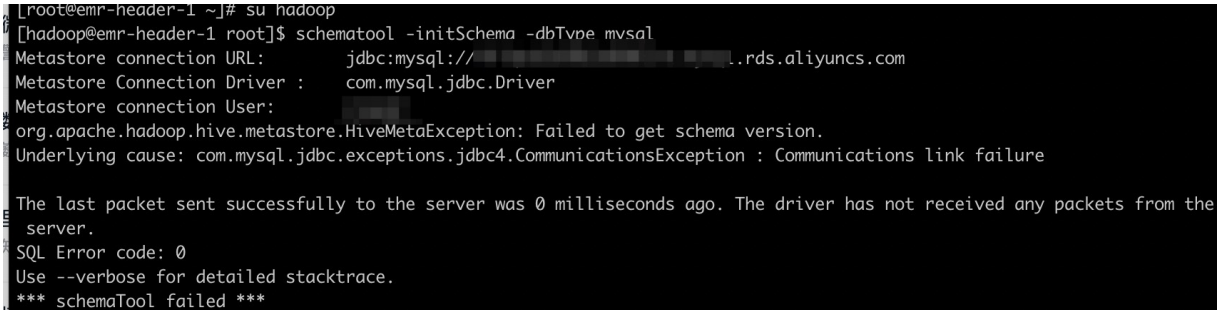
如果提示 `java.net.SocketTimeoutException` 异常信息，则您需要在HBase集群的安全组中新增安全组规则给Hive集群开放端口，示例如下。

内网入方向 | 内网出方向

授权策略	协议类型	端口范围	授权类型	授权对象	优先级	操作
允许	TCP	16000/16000	地址段访问		1	克隆 | 删除
允许	TCP	16020/16020	地址段访问		1	克隆 | 删除
允许	TCP	2181/2181	地址段访问		1	克隆 | 删除
允许	TCP	22/22	地址段访问		1	克隆 | 删除

Metastore初始化时提示Failed to get schema version异常信息，该如何处理？

异常详细信息如下图。



```
[root@emr-header-1 ~]# su hadoop
[hadoop@emr-header-1 root]$ schematool -initSchema -dbType mysql
Metastore connection URL:        jdbc:mysql://[redacted].rds.aliyuncs.com
Metastore Connection Driver :    com.mysql.jdbc.Driver
Metastore connection User:       [redacted]
org.apache.hadoop.hive.metastore.HiveMetaException: Failed to get schema version.
Underlying cause: com.mysql.jdbc.exceptions.jdbc4.CommunicationsException : Communications link failure

The last packet sent successfully to the server was 0 milliseconds ago. The driver has not received any packets from the
 server.
SQL Error code: 0
Use --verbose for detailed stacktrace.
*** schemaTool failed ***
```

请检查RDS MySQL安全组配置，确认RDS对EMR集群已开通安全组白名单。开通白名单详情，请参见设置IP白名单。

如果Hive元数据信息中包含中文信息，例如列注释和分区名等，该如何处理？

您可以在对应RDS数据库中逐条依次执行如下命令修改相应字段为UTF-8格式。

1. 执行以下命令，修改COMMENT列的数据类型。

   ```
   alter table COLUMNS_V2 modify column COMMENT varchar(256) character set utf8;
   ```

2. 执行以下命令，修改表TABLE_PARAMS中PARAM_VALUE列的数据类型。

   ```
   alter table TABLE_PARAMS modify column PARAM_VALUE varchar(4000) character set utf8;
   ```

3. 执行以下命令，修改表PARTITION_PARAMS中PARAM_VALUE列的数据类型。

   ```
   alter table PARTITION_PARAMS modify column PARAM_VALUE varchar(4000) character set utf8;
   ```

4. 执行以下命令，修改PKEY_COMMENT列的数据类型。

   ```
   alter table PARTITION_KEYS modify column PKEY_COMMENT varchar(4000) character set utf8;
   ```

5. 执行以下命令，修改表INDEX_PARAMS中PARAM_VALUE列的数据类型。

```
alter table INDEX_PARAMS modify column PARAM_VALUE varchar(4000) character set utf8;
```